Statistical Discrimination between Mammary Cancer and Mastopathy

Diana Lucía Londoño Londoño

Supervisor: Prof. Milan Stehlík

Advisor: Prof. Orietta Nicolis

This dissertation is submitted for the degree of

Doctor en Estadística

January 2019

Executive Summary

Breast cancer is currently considered a high frequency pathology in the world and one of the causes of higher mortality among women. According to recent statistical data published by the World Health Organisation (WHO), breast cancer accounts for 23% of all cancer related cases and 14% of all cancer related deaths among women worldwide (Zhou et al., 2017). Cancer is a group of diseases that cause cells in the body to change and grow out of control (Albertson and Pinkel, 2003). However, if breast anomalities are detected and diagnosis are made at early stages, studies show that the chances of survival can be greatly improved (Kestener et al., 2001; Martin et al., 1979). Currently the most reliable imaging technique for the detection of such anomalities is Mammography, or X-ray examination, that also plays an important role in control during and after the treatment. The early detection of breast cancer in asymptomatic women using breast screening mammography is currently the most effective imaging technique for the detection of such anomalities to reduce the morbidity and mortality associated with breast cancer (Lauby-Secretan et al., 2015). However, high imprecision and false / rate diagnostics makes problem hard.

Mammographies are low dose X-ray projections of the breast, and it is the best method for detecting cancer at the early stage. One of the stages present in the analysis of these mammography views involves the identification and classification of breast lesions, such as breast masses and micro-calcifications. Microcalcification is an effective indicator of early breast cancer (Wang et al., 2016).This identification and classification is usually performed manually by a radiologist, who should give a more accurate diagnosis. Furthermore, efficacy is often highly correlated with radiologist expertise and workload.

The interpretation of mammograms is however not an easy task. The mammographic appearance of normal tissue is highly variable and the radiological findings associated with breast cancer can be very complex. On the other hand, it has been estimated that $10-30\%$ of cancers which could have been detected are missed (Martin et al., 1979) and a high percent of the patients that are called back at screening turn out not to have cancer (Kestener et al., 2001). Reading a mammogram image is a skill that physicians develop over time, and confidently stating whether findings are cancerous or not is often quite difficult. Visual similarity between normal dense tissue and many types of breast cancers may also result in false negatives, and together with a significant false positive rate, this results in the diminished efficacy of screening mammography.

Discrimination between cancer tissue and mastopathy (benign and normal alterations mammary tissues) with an accurate probability is a crucial aid in supporting clinical decisions. Improving both the specificity and the sensitivity of mammographic diagnoses is an important goal in improving prognoses and reducing the number of unnecessary procedures or surgical operations. Dhungel et al. (2015) said that detection sensitivity can be improved in clinical practice with the use of an automated mass detection system that acts as a "second" opinion, which can reduce to some extent, the current dependence on the radiologist's experience and workload. Detection and characterization of breast lesions are important topics in computer-aided detection/diagnosis (CADe/CADx), which have been highly used in the field of medical image processing. Recently, different statistical methods have been proposed: stochastic geometry (Stehlík et al., 2012; Rezai-Rad and Jamarani, 2005), random carpet model (Hermann et al., 2015), percolation (Privman et al., 2016; Gorshkov et al., 2016; Stehlík et al., 2018), fractal and multifractal measures (Stehlík et al., 2012; Reljin and Reljin, 2002; Derado et al., 2008), wavelets (Nicolis et al., 2011; Kestener et al., 2001; Ramírez-Cobo and Vidakovic, 2013; Ramirez-Cobo et al., 2011), texture measures (Popescu et al., 2017; Dobrescu and Popescu, 2011; Khemis et al., 2016), machine learning (Yassin et al., 2017; Kourou et al., 2015) and aggregation (Mohamed et al., 2018; Gallego-Posada et al., 2016) of information of these methods to make decisions and to find a better model to describe this diseases in human pathology. Note also that this use of information contained in the background tissue of images is novel, since most of the references found in literature dealing with breast cancer detection methods are based on microcalcifications (Wang and Karayiannis, 1998; Netsch and Peitgen, 1999; Kestener et al., 2001; El-Naqa et al., 2002; Karssemeijer, 1992).

Here we use fractal analysis, multifractal analysis and texture analysis by obtaining relevant features to characterize the complex structure of the breast tissue. We compare this measures using stochastic geometry concepts, several techniques of image processing and automated statistical method. This classifying measures based on background tissue would be a new tool to be used in combination with existing clinical diagnostic tools, thus improving the power of non-invasive diagnostic techniques, which is the aim of this thesis. The novelty of this work is to analyze the efficency of the combining those techniques to differentiate between mastopathic versus cancerous tissue in the backgrounds of mammogram and histological images by using free databases as **MIAS** (Suckling et al., 1994) and **BreaKHis** (Spanhol et al., 2016).

This thesis is organized as follows:

Theory behind the mathematical concepts of fractals, multifractals, texture and percolation, primarily based upon works by Kenneth Falconer(Falconer, 1986, 1997, 2004), W. Steeb (Dekking and Meester, 1990) and B. Mandelbrot (Dekking and Meester, 1990; Lovejoy and Mandelbrot, 1985; B. Mandelbrot, 1983), will be described in chapters 1 , 2 and 3. Each section proposes a unified intuitive approach of the core concepts behind monofractals and multifractals, embedded in the field of measure theory.

Chapter 1 is devoted to the fractal theory and iterated function system that are needed for the generation of fractal sets: self-similar sets and the understanding of the multifractal theory. Since fractals and multifractals are based on "measures" or "mass distributions" this chapter begins with a section on measure theory. Fractal theory has been used in various applications to calculate the fractal dimension of an image which is a real number describing its structure or irregularity. Random and percolation fractals concepts are used for mathematical modeling of fractal sets, specifically, random Sierpiński carpet. Here a simulation study is presented in order to tissue growth modelling with the given percolation approach.

In Chapter 2 we look at the concept "fractal", different definitions of "fractal dimension" and some texture approaches. In this we introduce the methods used for calculating fractal and textural features measures. From the most simple box counting fractal method to multifractal gliding method. We extract statistical features from GLCM and we will classify mammogram images by using free databases (Suckling et al., 1994). Here we pretend to undertand the complexity of the structure of benign, malignant and normal breast cancer characterizing them through texture and fractal analysis.

The multifractal theory is addressed in chapter 3. Then two approaches to define the multifractal tool are presented: the coarse and the fine theory. Also, two proposed methods for calculating the multifractal exponents for a discrete image are presented and the results of this method for two real medical data are given: moment method or moment sums and wavelet-based multifractal spectrum. This teories are detailed, campared and applied to the two databases of images mentioned above. Multifractal theory can be considered as an extension of fractal theory and since some natural phenomena (including natural images) might be better described by the multifractal theory, it is interesting to look at the introduction of this tool to image analysis. The multifractal approaches will provide us with a "spectrum of fractal dimensions" characterizing the image and could potentially give us more information about the image compared to the single fractal dimension. Additionally,

a series of relevant parameters obtained from the wavelet-based multifractal spectrum are used for the characterization of breast tissue.

Finally, in Appendix A we introduce the code for generating random SC in the R package FractalParameterEstimation (Hermann et al., 2017). In appendix B we explain some of the mathematical subtleties allowing the wavelet methodology to be applied to calculate fractal dimension to 2D images.

To

Juan Camilo Herrera Patiño

Acknowledgements

Firstly, I would like to express my sincere gratitude to my supervisor and advisor, respectively, Prof. Milan Stehlík and Prof. Orietta Nicolis for the continuous support of my Ph.D study and related research, for his patience, motivation, and immense knowledge. His guidance helped me in all the time of research and writing of this thesis. I could not have imagined having a better advisers and mentors for my Ph.D study. I have to especially thank them, who trusted me and my potential, for the teachings, for the way that they treats their students, as equals, always with patience, availability and recognition.

My sincere thanks also goes to all the professors of the Institute of Statistics and Mathematics of the University of Valparaíso, who contributed to my training to reaching this goal. In particular, I am grateful to Prof. Rodrigo Salas and Prof. Rodrigo Castro for enlightening me with his vast knowledge in the development of this thesis. Without they precious support it would not be possible to conduct this research. I am profoundly grateful to all those who once helped me in a kind way, secretaries, assistants and many people I met during these years at this university.

Last but not the least, I would like to thank all my family: specially my parents, Elena and Ramiro, for giving birth to me at the first place and for the education that made possible for me to reach this point in my life, and for the understanding of my absence during these 4 years. To my brothers, sisters, nephews and nieces for supporting me spiritually throughout writing this thesis and my life in general.

A special thank to all my friends, that, through all these years, contributed to my professional growth, as well as on my personal growth. I am profoundly grateful with all them that on many times, even without knowing, turned bad moments into good moments during this work. To my friend Camilo, always present in my memory and whose energy and affection always urged me to keep going, this work is dedicated to him.

Thank you all.

Abstract

Spanish Version

El enfoque principal de esta tesis es la discriminacion entre tejidos mamarios cancerosos y tejdos mastopáticos, que son alteraciones benignas y malignas de los tejidos mamarios, lo anterior utilizando varios métodos estadísticos como geometría estocástica, modelos de carpetas aleatorias, wavelets, características de textura, análisis multifractal y agregación de información de estos métodos para tomar decisiones y encontrar el mejor modelo que describa estas enfermedades en la patología humana.

Para esta caracterización, proponemos un método computarizado para la clasificación automatizada del tejido mamario utilizando algunos parámetros estadísticos, como los índices fractales, parámetros multifractales, la lacunaridad, la succolaridad y las características de la textura clásica de la matrix de co-ocurrencia para analizar el tejido mamario. En esta tesis, proporcionamos un enfoque computable para la percolación, con posibles aplicaciones para el modelado del crecimiento de tejidos y se realiza un análisis multifractal sobre la base de un enfoque generalizado de dimensión fractal como también del enfoque multifractal del espectrum basado wavelet.

Analizamos el uso de los multifractales como una herramienta en la caracterización de la mama usando dos bases de datos gratuitas de imágenes de tejido tumoral (Suckling et al., 1994; Spanhol et al., 2016). Nuestro objetivo es explorar el comportamiento multifractal de diferentes tejidos mamarios complejos. El análisis multifractal es una forma útil de caracterizar sistemáticamente la heterogeneidad espacial de los patrones fractales teóricos y experimentales. En esta tesis, utilizamos dos algoritmos interesantes para el análisis multifractal de los tejidos mamarios: método de momentos o dimension fractal generalizada y un enfoque basado en wavelet.

Utilizaremos el espectro multifractal y sus descriptores en algoritmos de clasificación para discriminar el tejido mastopático del tejido canceroso. Se esperan diferencias significativas en la distribución subyacente entre los dos grupos.

English Version

The main focus of this thesis is the discrimination between tissues of cancer and mastopathy (benign and malignant alterations mammary tissues) using several statistical methods: stochastic geometry, random carpet model, wavelets, texture features, multifractal analysis and aggregation of information of these methods to make decisions and to find a better model to describe this diseases in human pathology.

For this characterization we proposes a computerized method for the automated classification of breast tissue using some statistical parameters such as the fractal indices, lacunarity, succolarity and classical texture features of GLCM to analyze breast tissue. In this thesis we provide computable approach to percolation, with possible applications to tissue growth modelling and a multifractal analysis on the basis of a generalized fractal dimension approach is conducted as well as the multifractal approach of the wavelet-based spectrum.

We look at the use of multifractals as a tool in image analysis on two free database of images of breast tumor tissue (Suckling et al., 1994; Spanhol et al., 2016). We aim to explore the multifractal behavior of different complex breast tissues. Multifractal analysis is a useful way to systematically characterize the spatial heterogeneity of both theoretical and experimental fractal patterns. In this thesis, we use two interesting algorithms for multifractal analysis of breast tissues: moment method or fractal dimension generalized and wavelet based approach.

We will use multifractal spectrum and its descriptors in classification algorithms for discriminating the mastopathic from the cancerous tissue. Significant differences in the underlying distribution between the two groups is expected.

General Objetive

The main focus of this thesis is the discrimination between tissues of cancer and mastopathy using several statistical methods and aggregation of information of these methods to make decisions and to find a better model to describe this diseases in human pathology.

The second aim is improve the model presented by Hermann et al. (2015) who introduced Random Carpet Model (RCM) which was very flexible model with a good fit to both masthopathy and mammary cancer. RCM has two parameters, p and q. The relation of p, q to percolation parameter can be substantial for understanding the real dimension of parametric space.

Specific Objective

1. We aim to improve the quality of statistical discrimination between masthopathy and mammary cancer tissues.

2. Computing/ developing the optimal designs for better estimation/prediction of parameters for Random Carpet Model (RCM).

3. Implementing computerized method for the automated classification using fractal, multifractal and texture features to analyze breast tissue.

Table of contents

List of figures

List of tables

Chapter 1

Geometric Properties of Statistically Self-similar Measures and Fractal Percolation

1.1 Introduction

Fractal geometry can be useful to describe the complex architecture of the tissue, the pathology of the tumors and can be applied for cancer detection. In mathematics, this concept has a long history and was introduced by Benoit Mandelbrot (B. Mandelbrot, 1983) based on his paper on self-similarity in which he discusssed the mathematical concept "*fractional dimensions*". By 1977, the mathematician was forced to give a formal definition that allowed to distinguish more clearly a fractal entity. In order to do so, he resorted to the old Hausdorff measure concept and in response to pragmatism defined, in general, all fractals as the set of forms with a fractional dimension. He introduced the term "fractal" (from the latin fractus, meaning "broken") to characterize spatial or temporal phenomena that are continuous but not differentiable. Unlike more familiar Euclidean constructs, every attempt to split a fractal into smaller pieces results in the resolution of more structure.

The two fundamental characteristics of fractal objects are:

1. Self-similarity: That is, those objects in which the smallest details that compose it have some statistical relationship with their global properties, repeating such details in an infinite way.

2. Fractal dimension or Hausdorff dimension: It is considered the main concept of Fractal Geometry, since fractal objects are characterized by having a fractional dimension.

Fractal objects and processes are therefore said to display "self-invariant" (self-similar or self-affine), scale independence, complexity and infinite length properties. Self-similar objects are isotropic upon rescaling, whereas rescaling of self-affine objects is directing-dependent (anisotropic). Thus the trace of particulate Brownian motion in two-dimensional space is self-similar, whereas a plot of the x-coordinate of the particle as a function of time is self-affine.

Self-similarity is an essential property of fractals in cancer and it is perhaps the main reason for the striking beauty of so many fractals. In fractal analysis, the Euclidean concept of "length" is viewed as a process and self-similarity also implies a scale-invariant property. This procedure may be quantified by the fractional dimension.

In Theiler (1990) offers a wide differente concepts between self-similarity and self-affinity. A set is strictly self-similar if it can be expressed as a union of sets, each of which is a reduced copy of (is geometrically similar to) the full set. However most fractal-looking objects in nature do not display quite this precise form. A fractal is self-affine if it can be decomposed into subsets that can be linearly mapped into the full figure. If this linear map involves only rotation, translation, and (isotropic) dilation, then the figure is self-similar. For a self-affine map, the contraction in one direction may differ from the contraction in another direction. The class of self-affine fractals therefore includes the class of self-similar fractals. The distinction between self-similar and self-affine fractals is not always made in practice. The concept is most useful in cases for which there are preferred global directions: in fractal surfaces, for example, or for fractal profiles [these are continuous nowhere-differentiable functions $f(x)$ for which the graph $(x, f(x))$ is a set of fractal dimension (Carter et al., 1988; Dubuc et al., 1989)].

There are several fractal dimensions described in the literature (see Steeb, 2014a, Chapter 8). The two most used are the *Hausdorff dimension* and *capacity*. The definition of the Hausdorff dimension is as follows. Let X be a subset of $\mathbb{R}^n$. A cover of X is a (possibly infinite) collection of balls, the union of which contains X. The diameter of a cover $\mathscr{A}$ is the maximum diameter of the balls in $\mathscr{A}$ For d, ε, we define

$$\alpha(d,\varepsilon) := \inf_{\mathscr{A}=\text{cover of } X \text{ diam}\,\mathscr{A}\leq\varepsilon} \sum_{A\in\mathscr{A}} (\text{diam}A)^d$$

and

$$\alpha(d) := \lim_{\varepsilon \to 0} \alpha(d, \varepsilon)$$

It can be shown that there is a unique d_0 such that

$$d < d_0 \Rightarrow \alpha(d) = \infty$$

$$d > d_0 \Rightarrow \alpha(d) = 0$$

This d_0 is defined to be the Hausdorff dimension of X, Written HD(X).

The capacity C is a special case of the Hausdorff dimension (Steeb, 2014a) (also called box-counting dimension). In this case the covers consist of balls of uniform size ε. We put a grid on $\mathbb{R}^n$ and count the number of boxes that intersect the set X. Suppose we have such an ε-grid. Let $\mathbb{N}(\varepsilon) =$ the number of boxes that intersect the set X. Then $\mathbb{N}(\varepsilon \cdot \varepsilon^d)$ would seem to be a simplified version of $\alpha(d, \varepsilon)$ and the desired d can sometimes be computed as

$$\mathrm{C} := \lim_{\varepsilon \to 0} \frac{\log N(\varepsilon)}{\log \frac{1}{\varepsilon}}$$

This number C is called capacity. We have the inequality $\mathrm{C}(X) < HD(X)$.

Fractals reveal how irregular objects can be interpreted and described according to morphological characteristics. The irregular behaviors of these complex structures are difficult or impossible to quantify by standard modeling techniques or by traditional Euclidean geometry. However, to approximate fractals to cancer fits well as a morphometric tool to classify mammopathic tissues of cancerous tissues in diagnostic and prognostic processes.

The fractal analysis is an important tool for the diagnosis of breast cancer since it can quantify the irregularity and complexity of objects. Self-similarity is an essential property of fractals in cancer and it is perhaps the main reason for the striking beauty of so many fractals. Self-similarity also implies a scale-invariant property and it may be quantified by a fractal dimension (FD), which we will see in greater depth in Chapter 2.

Excellent summaries of basic concepts of fractal geometry can be found in the books (B. Mandelbrot, 1983; Falconer, 1986, 1997, 2004).

Definition 1. *A fractal is by definition a set for which the Hausdorff dimension strictly exceeds the topological dimension.*

In this chapter we will study fractals from an optics associated with exact self-similarity, we will define the concept of iterated function systems (IFS) and using techniques of functional analysis and measure theory, we will demonstrate the existence of these objects called Fractals. Also, as a application we provide computable approach to percolation, with possible applications to tissue growth modelling Stehlík et al. (2018).

1.2 Hausdorff Measure

Measures have a central place in fractal geometry. They are a major tool in the mathematics of fractals, but also, measures may exhibit fractal features which may be studied in their own right. Basically in Falconer (1997), a measure is a way of ascribing a numerical size to sets so that the principle "the whole is the sum of the parts" applies. If a set is decomposed into a finite or countable number of pieces in a reasonable way then the measure of the whole set is the sum of the measures of the pieces.

The idea of defining measures using covers of sets was introduced by Carathéodory (1914). Hausdorff (1919) used this method to define the measures that now bear his name, and showed that the middle third Cantor set has positive and finite measure of dimension $\log 2/\log 3$. Properties of Hausdorff measures have been developed ever since, not least by Besicovitch and his students.

Technical aspects of Hausdorff measures are discussed in more details in Falconer (1986, 2004).

Before we begin to analyze fractals, we must first introduce some concepts in measure theory. These tools will prove useful for studying the exotic nature of fractals. We will need to use a measure other than the usual Lebesgue measure. We recall the Lebesgue measure here:

Definition 2. *Let $\{B_i\}$ be a countable collection of $n-$dimensional boxes, i.e., subsets of $\mathbb{R}^n$ such that $B_i = \{(x_1,\ldots,x_n) \in \mathbb{R}^n : a_j \leq x_j \leq b_j\}$ for $a_j < b_j$ and $j = \{1,\ldots,n\}$. We will write the $n-$dimensional volume of B_i as $\mathrm{vol}(B_i) = (b_1-a_1)(b_2-a_2)\cdot(b_n-a_n)$. Then the Lebesgue measure of a subset E of $\mathbb{R}^n$ is*

$$\mathscr{L}^n(E) = \inf\left\{\sum_i \mathrm{vol}\,(B_i) : E \subset \bigcup_i B_i\right\}$$

Our intuitive understanding of length, area, and 3-dimensional volume is given by the Lebesgue measure, in particular, $\mathscr{L}^1$, $\mathscr{L}^2$, and $\mathscr{L}^3$, respectively. A generalization of Lebesgue measure, known as Hausdorff measure, is given below.

Definition 3. *For any subset U of $\mathbb{R}^n$, the diameter of U is*

$$\operatorname{diam} U = \sup\{|x-y| : x,y \in U\}$$

Recall that if U is any non-empty subset of $n-$dimensional Euclidean space, $\mathbb{R}^n$, the diameter of U is defined as $|U| = \sup\{|x-y| : x,y \in U\}$, i.e. the greatest distance apart of any pair of points in U.

Definition 4. *If $\{U_i\}$ is a countable (or finite) collection of sets of diameter at most δ that cover F, i.e. $F \subset \bigcup_{i=1}^{\infty} U_i$ with $0 \leq |U_i| \leq \delta$ for each i, we say that $\{U_i\}$ is a $\delta-$cover of F. Suppose that F is a subset of $\mathbb{R}^n$ and s is a non-negative number. For any $\delta > 0$ we define the $s-$dimensional Hausdorff measure of F*

$$\mathscr{H}^s_\delta(F) = \inf\left\{\sum_{i=1}^{\infty} |U_i|^s : \{U_i\} \text{ is a } \delta - \text{cover of } F\right\} \tag{1.1}$$

This is similar to the Lebesgue measure. Both make use of a countable collection of sets that cover the set we wish to measure, and then calculate its measure from the "tightest" covering. Thus we look at all covers of F by sets of diameter at most δ and seek to minimize the sum of the sth powers of the diameters. As δ decreases, the class of permissible covers of F in Equation (1.1) is reduced. Therefore, the infimum $\mathscr{H}^s_\delta(F)$ increases, and so approaches a limit as $\delta \to 0$. We write

$$\mathscr{H}^s = \lim_{\delta \to 0} \mathscr{H}^s_\delta(F). \tag{1.2}$$

This limit exists for any subset F of $\mathbb{R}^n$, though the limiting value can be (and usually is) 0 or ∞. We call $\mathscr{H}^s_\delta(F)$ the *s-dimensional Hausdorff measure of F*.

To finish establishing the definition of Hausdorff measure, we will prove that it is an *outer measure, Borel sets and support of a measure*. (Note that it is a fact from measure theory that an outer measure can become a measure if restricted to measurable sets.)

Definition 5. *A function μ is an outer measure if it is a nonnegative function defined on all subsets of $\mathbb{R}^n$ such that*

- $\mu(\varnothing) = 0$,
- *monotonicity:* $\mu(A) \leq \mu(B)$ *if* $A \subseteq B$, *and*
- *countable subadditivity:* $\mu\left(F \subset \bigcup_{i=1}^{\infty} A_i\right) \leq \sum_{i=1}^{\infty} \mu\left(A_i\right)$ *for a countable (or finite) sequence of sets* $\{A_i\}$.

Countable subadditivity says that if a set is decomposed into smaller parts (which may overlap), the sum of the measure of the parts is at least equal to the measure of the whole set. The second property says that the function is monotone, i.e. the larger the set, the larger the measure, that is a natural property for a measure not implied by the third property.

The collection of sets that can be measured by an outer measure μ forms a $\sigma-$algebra and by restricting μ to this $\sigma-$algebra we will get a measure.

The power set of X is a $\sigma-$algebra. Any collection S of subsets of X can always be contained in the power set and therefore in a $\sigma-$algebra. If we take the intersection of all possible$\sigma-$algebras that contains S we get the smallest $\sigma-$algebra containing S and we call this the $\sigma-$algebra generated by S. For proof see (Edgar, 2007) p.133.

The following definition defines a measurable collection of subsets of X including all sets that will be interesting to look at in any practical situation.

Definition 6. *The $\sigma-$algebra generated by the collection of all open (or equivalently closed) subsets of X is called the Borel algebra on X, denoted $\beta(X)$. The sets that belong to $\beta(X)$ are called the Borel sets with respect to$X$.*

We know from Definition 6 that X and its complement $\varnothing$ are measurable sets, but we can also measure any set that can be constructed using a sequence of countable unions or intersections with the open or closed sets, i.e. the Borel sets (see (Falconer, 1986) or (Falconer, 2004)). The Borel algebra is the smallest collection of subsets of X with the following properties:

1. Every open and closed set is a Borel set.
2. The union of every finite or countable collection of Borel sets is a Borel set, and the intersection of every finite or countable collection of Borel sets is a Borel set.

All the subsets of $\mathbb{R}^n$ that we will work on later will be Borel sets. The multifractal measures will be outer measures for which the Borel sets are measurable.

Definition 7. *A measure μ, or the restriction of an outer measure to its measurable sets, is called a Borel measure on X if the Borel sets with respect to X can be measured by μ.*

We will mainly consider finite Borel measures which means that the set function only takes on values less than infinity. Most of the time we will also talk about Borel probability (or normalized) measures, that is Borel measures where $\mu(X) = 1$. Such a measure can be constructed from any finite Borel measure (with $\mu(X) \neq 0$) by

$$P(A) = \mu(A)/c$$

for some A in the domain of μ where c is the total measure $\mu(X) \in (0, \infty)$. To prove that $P(A)$ is a measure it is necessary verify the properties given in definition 5.

The support of a measure tells us where in X the measure takes place, that is where it is non-zero. The support is defined as the largest closed subset of X for which every open neighbourhood of every point in the set has positive measure.

Definition 8. *The support of a measure μ on X is*

$$spt(\mu) = X \setminus \bigcup_{\mu(F)=0} F \tag{1.3}$$

where F are open sets with zero measure.

Theorem 1. *The Hausdorff measure is an outer measure (Falconer, 2004).*

With the definition of Hausdorff measure sufficiently established, we will proceed to study its properties and behavior. Indeed, it has many nice properties. The following properties of the Hausdorff $s-$dimensional measure are essential in fractal theory. We refer to the book of Falconer (2004) for a detailed description and proof of this ones.

Theorem 2. *Let S be a similarity, i.e., a mapping $S : \mathbb{R}^n \to \mathbb{R}^n$ such that $|S(x) - S(y)| = \lambda |x - y|$ for all x, y in $\mathbb{R}^n$, with $\lambda, c, r > 0$. For a subset of $\mathbb{R}^n$ the $s-$dimensional measure of Hausdorff has the following properties:*

1. *If S is a isometry, then $\mathscr{H}^s(S(A)) = \mathscr{H}^s(A)$.*
2. *If S is a λ ratio homothety then $\mathscr{H}^s(S(A)) = \lambda^s \mathscr{H}^s(A)$.*
3. *If S is a Lipschitz mapping with Lipschitz constant c, then $\mathscr{H}^s(S(A)) \leq c^s \mathscr{H}^s(A)$.*

4. *If S is a similarity, such that $|S(x)-S(y)| = r|x-y|$, with $r \neq 0$, then $\mathscr{H}^s(S(A)) = r^s \mathscr{H}^s(A)$.*

1.3 Iterated Function System

Barnsley and Demko (1985) said that J. E. Hutchinson, in 1981 was the first mathematician who, studying the common properties (compactness, self-similarity, etc.) of the fractals already known, elaborated a unified theory for the obtaining of a wide class of fractal sets: self-similar sets. Generalizing the method of Hutchinson, in 1985, M. F. Barnsley uses the systems of iterated functions for the generation of fractal sets. Iterated Function System (IFS) provide a very convenient way of representing and reconstructing many fractals that in some way are made up of small images of themselves. To define an IFS we need the definition of a complete metric space.

1.3.1 Metric spaces

We are mainly interested in random fractals and in the description of a specific fractal structure such as Sierpinski carpet, as well as also some properties such as kronecker product, generalization of Mandelbrot's percolation process and Frobenius norm. Nevertheless, we must make another detour before analyzing those approaches of fractals. To guarantee the existence and uniqueness of fractals as produced by contracting similarities, we will need some tools from the theory of metric spaces. This will allow us to more rigorously define contractions, and it will give us the necessary fixed point theorem.

Consider the metric space $(\mathbb{R}^n, d)$; suppose that d is the Euclidean metric, we know that $\mathbb{R}^n$ with this metric is complete, that is, all Cauchy's sequences are convergent, remember that:

$$\{x_n\}_{n\in\mathbb{N}} \forall \varepsilon > 0, \exists N \in \mathbb{N} \quad \text{such that} \quad d(x_k, x_r) < \varepsilon \quad \text{always that} \quad k, r \geq N$$

that is, the elements of the succession's tail are arbitrarily close to each other.

Not all metric space is complete, for example, $\mathbb{Q}$ with the Euclidean metric is not complete, since it is easy to find a succession of rational numbers that converge to an irrational one.

Definition 9. *A pair (X,d), where $d : X \times X \to R$, is a metric space if d satisfies*

- $d(x,y) \geq 0$, *with equality only when* $x = y$,
- *If* $x \neq y$,*then* $d(x,y) > 0$
- $d(x,y) = d(y,x)$ *symmetry, and*
- $d(x,y) \leq d(x,z) + d(z,y)$ *for all z in X*

The nonnegative real number $d(x,y)$ is called the distance from x to y. Then d is called a metric. If every Cauchy sequence in X converges, X is a complete metric space.

Now we will study the space in which some of our fractals live, the metric space of compact figures, herein we must remember a couple of concepts of functional analysis. For example, the *closed ball* of centre x and radius r is defined by $B(x,r) = \{y : |y-x| \leq r\}$. Similarly the *open ball* is $B^{\circ}(x,r) = \{y : |y-x| < r\}$. Thus the closed ball contains its bounding sphere, but the open ball does not. Of course in $\mathbb{R}^2$ a ball is a disc and in $\mathbb{R}^1$ a ball is just an interval.

Definition 10. *1. From time to time we refer to the* $\delta-$*neighbourhood or* $\delta-$ *parallel body,* A_δ*, of a set A, that is the set of points within distance* δ *of A; thus* $A_\delta = \{x : |x-y| \leq d\}$ *for some y in A; see figure 1.1.*

$$A^{\delta} = \bigcup_{x \in A} B(x,\delta)$$

where $B(x,\delta)$ *is the closed ball of centre x and radius* δ*.*

2. *Let* A, B *be two compact subsets of* $\mathbb{R}^n$*, we define the Hausdorff distance between A and B as*

$$d_H(A,B) = \min_{\delta > 0} \left\{ \delta \geq 0 : A \subseteq B^{\delta} \ y \ B \subseteq A^{\delta} \right\}$$

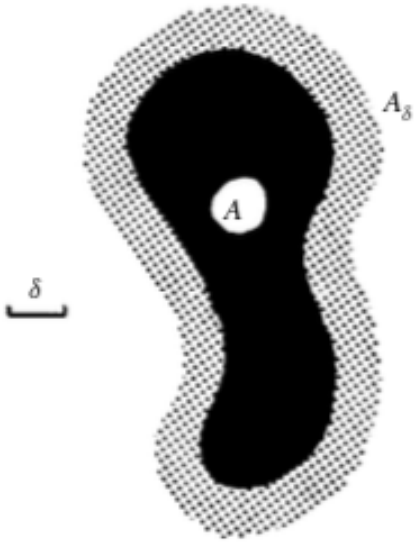


Fig. 1.1 A set A and its $\delta-$neighbourhood A_δ

It is easy to prove that effectively d_H is a distance in space from the compact subsets of $\mathbb{R}^n$; in addition to

$$d_H(A,B) = \max\left\{\max\left\{d\left(A,b\right) : b \in B\right\}, \max\left\{d\left(a,B\right) : a \in A\right\}\right\}$$

A simple check shows that d is a metric or distance function, that is, satisfies the three requirements (i) $d(A,B) \geqslant 0$ with equality if and only if $A = B$, (ii) $d(A,B) = d(B,A)$, (iii) $d(A,B) \leqslant d(A,C) + d(C,B)$ for all $A,B,C \in S$. The metric d is known as the *Hausdorff metric* on S. In particular, if $d(A,B)$ is small, then A and B are close to each other as sets.

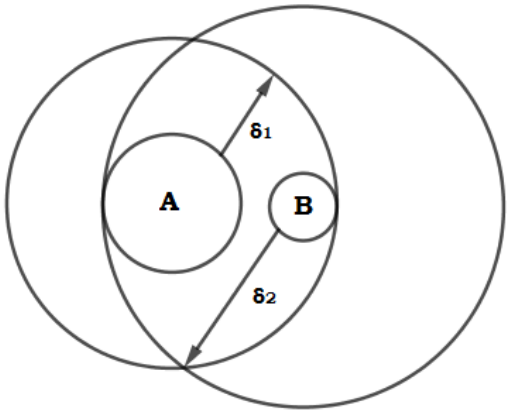


Fig. 1.2 Hausdorff distance between convex polygons A and B, $d(A,B) = \delta_2$

Let D be a closed subset of $\mathbb{R}^n$. For a subset E of D, write $N_\delta(E)$ for the $\delta-$neighborhood of E, i.e,

$$N_\delta(E) = \{x \in D : |x-a| < \delta \quad \text{for some } a \in E\}.$$

Then let Ω be the set of all nonempty compact subsets of D. We associate the following metric to Ω:

Definition 11. *The Hausdorff metric on Ω for any A, $B \in \Omega$ is*

$$d_H(A,B) = \inf\{\delta : A \subset N_\delta(B) \quad \textit{and} \quad B \subset N_\delta(A)\}.$$

We work in non-empty closed subset X of $\mathbb{R}^n$,

Theorem 3. *The set Ω associated with the Hausdorff metric d_H is a complete metric space (see (Steeb, 2014a)).*

Definition 12. 1. *Given a metric space X with metric d, a mapping $S : X \to X$ is a contraction if*

$$d(S(x),S(y)) \leq c\, d(x,y)$$

for some constant positive $c < 1$ and for all $x,y \in X$. If equality holds everywhere, we call S a contracting similarity.

2. *x_0 is a fixed point of mapping $S : X \to X$ if $S(x_0) = x_0$.*

The following theorem will play a central role in our construction of fractals. We refer to the book of Falconer (1997) for a thorough understanding of this theorem.

Theorem 4 (Banach's Contraction Mapping Theorem)**.** *If X is a complete metric space, and if S is a contraction of X into X, then there exists one and only one x_0 in X such that $S(x_0) = x_0$.*

Given the importance of demonstrating this theorem for the construction of fractals, we will make an outline of the demonstration.

Proof. Let $x_0 \in X$ be an arbitrary point. Recursively define the sequence x_n such that $x_{n+1} = S(x_n)$ i.e.

$$x_0, \quad x_1 = S(x_0), \quad x_2 = S(x_1), \quad \ldots, \quad x_{n+1} = S(x_n);$$

Then

$$d(x_{n+1}, x_n) = d(S(x_n), S(x_{n-1})) \leq c \;\; d(x_n, x_{n-1}) \leq c^n d(x_1, x_0).$$

Let's see that x_n is a Cauchy sequence, in effect,

$$\begin{aligned} d(x_{m+1}, x_m) &= d(S(x_m), S(x_{m-1})) \\ &\leq c d(x_m, x_{m-1}) \\ &= c d(S(x_{m-1}), S(x_{m-2})) \\ &\leq c^2 d(x_{m-1}, x_{m-2}) \\ &\cdot \\ &\cdot \\ &\cdot \\ &\leq c^m d(x_1, x_0) \end{aligned}$$

For $n < m$, it follows from the triangle inequality that

$$\begin{aligned} d(x_n, x_m) &\leq \sum_{i=n+1}^{m} d(x_i, x_{i-1}) \\ &\leq \sum_{i=n+1}^{m} c^{i-1} d(x_1, x_0) \end{aligned}$$

since $0 < c < 1$; then $1 - c^{n-m} < 1$; thus

$$\leq \sum_{i=n+1}^{m} \frac{c^n}{1-c} d(x_1, x_0)$$

Since $c < 1$, $d(x_n, x_m)$ approaches 0 for large n, and so x_n is a Cauchy sequence. Since X is complete,x_n converges to some $x = \lim_{n\to\infty} x_n$ in X. Finally, from the triangle inequality,

$$\begin{aligned} d(x, S(x)) &\leq d(x, x_m) + d(x_m, S(x)) \\ &\leq d(x, x_m) + cd(x_{m-1}, x) \end{aligned}$$

Note that S is continuous since it is a contraction. Therefore,

$$S(x) = S\left(\lim_{n\to\infty} x_n\right) = \lim_{n\to\infty} S(x_n) = \lim_{n\to\infty} x_{n+1} = x,$$

and so x is indeed a fixed point.

$$S(x) = x$$

To prove the uniqueness of x, suppose that if $S(x_1) = x_1$ and $S(x_2) = x_2$, then

$$d(x_1, x_2) = d(S(x_1), S(x_2)) \leq cd(x_1, x_2)$$

Since $S(x_1) = x_1$ and $S(x_2) = x_2$, we have $d(x_1, x_2) \leq cd(x_1, x_2)$. This implies that $d(x_1, x_2) = 0$, or $x_1 = x_2$.

□

The following is the central theorem (Falconer, 2004; Steeb, 2014b), which will allow us to build fractals.

Theorem 5. *If*

$$\mathscr{H}^s(\mathbb{R}^n) = \{A \subset \mathbb{R}^n : A \text{ is compact}\},$$

then $(\mathscr{H}(\mathbb{R}^n), d_h)$ *is a complete metric space, also,* $\{A_n\}_{n\in\mathbb{N}}$ *is a Cauchy sequence in* $\mathscr{H}(\mathbb{R}^n)$ *and* $A_n \to A$*, then*

$$A = \{x \in \mathbb{R}^n : \exists\, x_m \to x \ \textit{with}\ x_m \in A_m\},\ \forall m \in \mathbb{N}$$

We will not make the proof of this theorem, however, given its importance we will explain the steps to follow in this one

1. Initially prove that $A \neq \emptyset$, for this to demonstrate that there is a Cauchy sequence $\{x_n\}_{n\in\mathbb{N}}$ with $x_n \in A_n \in \mathscr{H}(\mathbb{R}^n), \forall n \in \mathbb{N}$, and the limit of that sequence is an element of A.

2. Next show that A is closed and also that $\forall\delta,\ \delta > 0,\ \exists N \in \mathbb{N}$ such that $A \subseteq A_n^{\delta}$ if $n \geq N$.

3. A is bounded. This property is immediate and therefore, $A \in \mathscr{H}(\mathbb{R}^n)$.

4. Finally it is necessary to prove that $A_n \to A$, for this, use the fact that: $d_H(A, A_n) < \delta \to A_n \subseteq A^{\delta}$.

1.3.2 Iterated Function Systems

Many fractals are made up of parts that are, in some way, similar to the whole. For example, geometric structures such as the Sierpinski gasket, the Sierpinski carpet and the Koch curve are examples of exactly self-similar fractals or regular fractals, which exhibit self-similarity over an infinite range of length scales. Figure 1.3 displays the exactly self-similar fractals, Sierpinski carpets of fractal. These self-similarities are not only properties of the fractals: they may actually be used to define them. Iterated function systems do this in a unified way and, moreover, often lead to a simple way of finding dimensions.

The fractals we are interested in studying in this chapter are all self-similar, i.e., they are made up of several contracted copies of themselves. Clearly, fractals are fixed points under these contractions. We will use the concept of iterated function systems, under which fixed points are called attractors, to formalize this method of constructing fractals.

Definition 13. *An iterated function system (abbreviated "IFS") is a collection finite of contractions* $S_1, S_2, \ldots, S_m$*, with* $m \geq 2$*, on a closed subset* D *of* $\mathbb{R}^n$*, often* $D = \mathbb{R}^n$*.A mapping*

$S : D \to D$ is called a contraction on D if there is a number c with $0 < c < 1$ such that $|S(x) - S(y)| \leqslant c|x-y|$ for all $x, y \in D$. Clearly any contraction is continuous.

If equality holds, i.e. if $|S(x) - S(y)| = c|x-y|$, then S transforms sets into geometrically similar sets, and we call S a contracting similarity.

A nonempty compact subset F of D, is called an attractor (or invariant set) of the IFS if

$$F = \bigcup_{i=1}^{m} S_i(F)$$

The fundamental property of an iterated function system is that it determines a unique attractor, which is usually a fractal. To prove this, it is necessary to apply Theorem 4 to the contraction $S(\cdot) = \bigcup_{i=1}^{m} S_i(\cdot)$ on the complete metric space Ω.

Falconer (2004) gives two proofs of the fundamental result on IFSs. The first depends on Banach's contraction mapping theorem, and the second is direct and elementary.

Theorem 6. *Consider the iterated function system given by the contractions $S_1, \ldots, S_m$ on $D \subset \mathbb{R}^n$, so that*

$$|S_i(x) - S_i(y)| \leqslant c_i\,|x-y|\,(x,y) \in D$$

with $c_i < 1$for each i. Then there is a unique attractor F, i.e. a non-empty compact set such that

$$F = \bigcup_{i=1}^{m} S_i(F)$$

Moreover, if we define a transformation S on the class S of non-empty compact sets by

$$S(E) = \bigcup_{i=1}^{m} S_i(E)$$

for $E \in S$, and write S^k for the kth iterate of S (so $S^0(E) = E$ and $S^k(E) = S\left(S^{k-1}(E)\right)$ for $k \geqslant 1$, then

$$F = \bigcap_{k=0}^{\infty} S^k(E)$$

for every set $E \in S$ such that $S_i(E) \subset E$ for all i.

This theorem holds for all contractions, not only contracting similarities. We will give below an example of calculating the dimension of self-similar fractals, for the Sierpinski Carpet.

1.3.3 Sierpinski carpet

Sierpinski carpet (SC) is a fractal, which was introduced by Waclaw Sierpinski around 1920, as example of a curve in which every point is branching. This fractal can have different structures and is built upon a special algorithm. Starting from a square as basis (Fig. 1.3) and using an analogous process to the construction of the Sierpinski Triangle, we can build the so-called Sierpinski carpet. In the first step we start from the unit square and we eliminate the central square from side $1/3$ and to each of the 9 remaining squares we eliminate a square from side $(1/3)^2$ and so on. This mathematical model can then be used to model porous bodies, like cancerous tissue (Hermann et al., 2015). Fractals reveal how irregular objects can be interpreted and described according to morphological characteristics. The irregular behaviors of these complex structures are difficult or impossible to quantify by standard modeling techniques or by traditional Euclidean geometry. However, to approximate fractals to cancer fits well as a morphometric tool to classify mammopathic tissues of cancerous tissues in diagnostic and prognostic processes. The fractal analysis is an important tool for the diagnosis of breast cancer since it can quantify the irregularity and complexity of objects. Here lies the potential of fractal analysis as an important model of the irregular structures of tumor growth.

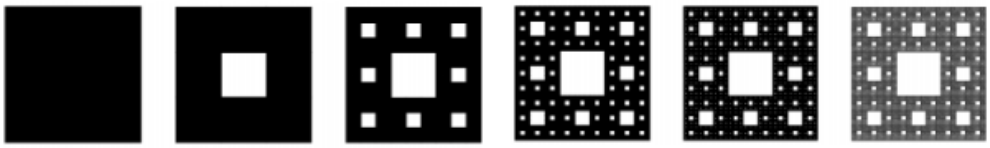


Fig. 1.3 Sierpinski Carpet by boxes of decreasing size

Definition 14. *The basic SC is a fractal subset of* $[0,1]^2$ *defined in a similar way to the classical Cantor set, except that one removes the middle square out of a* 3×3 *block. It can be regarded as a model space for studying diffusion in irregular media, with irregularities at many different length scales.*

One can define the basic SC in d dimensions, by starting with $[0,1]^d$*, dividing each cube into* 3^d *subcubes and removing the middle cube.*

1.3.4 IFS for SC

We will then give the IFS that gives rise to the SC. By definition 14, a SC first start with a solid (filled) square $\mathbf{E_0}$ $[0,1]^2$ and divide this into 9 smaller congruent squares. Remove

the interior of the center square (that is, do not remove the boundary) to get $\mathbf{E_1}$. Now subdivide each of the eight remaining solid squares into 9 congruent squares and remove the center square from each to obtain $\mathbf{E_2}$. Continue to repeat the construction to obtain a decreasing sequence of sets.

Intuitively, we consider a construction $F = \bigcap_{k=1}^{\infty} E_k$ where $[0,1]^2 = E_0 \supset E_1 \supset \ldots$ is a decreasing sequence of closed sets. The Sierpinski carpet is the intersection of all the sets in this sequence, that is, the set of points that remain after this construction is iterated infinitely often.

The original square is scaled by a factor $r = 1/3$. This is done 8 times followed by the necessary translations to arrange the eight squares as described Figure 1.3. If we take the original square to be a unit square with opposite corners at $(0,0)$ and $(1,1)$, then the IFS would be given by the following functions.

$$
\begin{aligned}
f_1(\mathbf{x}) &= \begin{bmatrix} 1/3 & 0 \\ 0 & 1/3 \end{bmatrix} \mathbf{x} \\
f_2(\mathbf{x}) &= \begin{bmatrix} 1/3 & 0 \\ 0 & 1/3 \end{bmatrix} \mathbf{x} + \begin{bmatrix} 0 \\ 1/3 \end{bmatrix} \\
f_3(\mathbf{x}) &= \begin{bmatrix} 1/3 & 0 \\ 0 & 1/3 \end{bmatrix} \mathbf{x} + \begin{bmatrix} 0 \\ 2/3 \end{bmatrix} \\
f_4(\mathbf{x}) &= \begin{bmatrix} 1/3 & 0 \\ 0 & 1/3 \end{bmatrix} \mathbf{x} + \begin{bmatrix} 1/3 \\ 0 \end{bmatrix} \\
f_5(\mathbf{x}) &= \begin{bmatrix} 1/3 & 0 \\ 0 & 1/3 \end{bmatrix} \mathbf{x} + \begin{bmatrix} 1/3 \\ 2/3 \end{bmatrix} \\
f_6(\mathbf{x}) &= \begin{bmatrix} 1/3 & 0 \\ 0 & 1/3 \end{bmatrix} \mathbf{x} + \begin{bmatrix} 2/3 \\ 0 \end{bmatrix} \\
f_7(\mathbf{x}) &= \begin{bmatrix} 1/3 & 0 \\ 0 & 1/3 \end{bmatrix} \mathbf{x} + \begin{bmatrix} 2/3 \\ 1/3 \end{bmatrix} \\
f_8(\mathbf{x}) &= \begin{bmatrix} 1/3 & 0 \\ 0 & 1/3 \end{bmatrix} \mathbf{x} + \begin{bmatrix} 2/3 \\ 2/3 \end{bmatrix}
\end{aligned}
$$

The Sierpinski carpet consists of eight self-similar pieces corresponding to the eight functions in the IFS.

1.4 Random Fractals

The term fractal usually refers to sets which, in some sense, have a self-similar structure. Already in the seventies of the last century Mandelbrot (B. Mandelbrot, 1983) made a compelling case for the importance of this concept in mathematical modelling. Indeed, some form of self-similarity is common in random sets, in particular those arising from stochastic processes. Therefore studying fractal aspects is an important feature of modern stochastic geometry.

In previous section, we saw that the fixed points of IFS are often referred to sets with obvious self-similarity. These are toy examples, tailor-made to study self-similarity in its tidiest form and are considered early progress in fractal geometry. An overview of the achievements in this period can be obtained from Falconer (2004) and Natoli (2012). However, *random fractals* do not have the self-similarity of their non-random counterparts, their non-uniform appearance is often rather closer to natural phenomena such as coastlines, topographical surfaces or cloud boundaries. Indeed, random fractal constructions are the basis of many impressive computer-drawn landscapes or skyscapes.

Most fractals discussed in this work involve a sequence of approximations E_k, each obtained from its predecessor by modification in increasingly fine detail, with a fractal F as a limiting set. A random fractal worthy of the name should display randomness at all scales, so it is appropriate to introduce a random element at each stage of the construction. By relating the size of the random variations to the scale, we can arrange for the fractal to be statistically self-similar in the sense that enlargements of small parts have the same statistical distribution as the whole set. This compares with (non-random) self-similar sets where enlargements of small parts are identical to the whole.

In the last ten years interest in this area has increased considerably, powerful techniques have been developed, and very substantial progress has been made. Typical examples of the fractals studied today are level sets of stochastic processes, the double points of random curves, or the boundary of excursions of random fields. The self-similar nature of these examples is typically less tidy and exploiting it means entering deep into the geometry of the sets.

Very roughly speaking, a set is self-similar if it can be decomposed into parts which look like scaled copies of the original set. This definition becomes particularly powerful when "look like" is interpreted in a statistical sense, i.e. if it can be decomposed into parts which have (up to scaling) the same distribution as the whole set.

There are two major aspects of the application of random fractals: the first is the modeling and characterization of natural forms and processes; while the second concerns the practical issue of computer simulation of such phenomena. In the first case one is interested in a physically based, accurate model that reflects all relevant aspects of the phenomenon and is in good agreement with theory and experiment. The goal is to improve understanding of the natural phenomenon in order to make predictions possible. Such modeling and analysis is of prime concern to many scientific areas dealing directly with nature (physics, chemistry, biology, geology, astronomy and so on).

For a better understanding of the topics that we will develop next, we refer to our paper (Stehlík et al., 2018) as a theoretical and practical result of the concepts presented here.

1.4.1 A random Sierpiński carpet

As we previously saw in 14, the Sierpiński carpet (SC) is the fractal, for which the construction begins with a square $[0,1]^2$. We can equivalently construct this fractal by using string rewriting beginning with a cell 1 and iterating the rules

$$\left\{ 0 \to \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \; 1 \to \begin{bmatrix} 1 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 1 \end{bmatrix} \right\}.$$

Here cell 0 indicates absence and cell 1 indicates presence of generated square. Random variation of SC can be defined by using of iterating the rules

$$\left\{ 0 \to \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \; 1 \to \begin{bmatrix} \mathscr{B}_{1,p_n} & \mathscr{B}_{2,p_n} & \mathscr{B}_{3,p_n} \\ \mathscr{B}_{4,p_n} & 0 & \mathscr{B}_{5,p_n} \\ \mathscr{B}_{6,p_n} & \mathscr{B}_{7,p_n} & \mathscr{B}_{8,p_n} \end{bmatrix} \right\}. \tag{1.4}$$

beginning again with a cell 1, whereas $\mathscr{B}_{i,p_n}, i = 1,\ldots,8$ are mutually independent random variables generated in n-th step from a Bernoulli distribution. Notice that "approximating" $[p,p,p,q]$ model, Hermann et al. (2015), is involved here. This notation naturally gener-

alizes this approximation into $[p,\dots,p,q]$ model and also $[p_1,\dots,p_{n-1},p_n]$ model. For illustration see Figure 1.4 and Figure 1.5.

Remark 1. *Naturally this can be generalized also in the following sense:*

$$1 \to \begin{bmatrix} \mathscr{B}_{p_{13,n}} & \mathscr{B}_{p_{23,n}} & \mathscr{B}_{p_{33,n}} \\ \mathscr{B}_{p_{12,n}} & \mathscr{B}_{p_{22,n}} & \mathscr{B}_{p_{32,n}} \\ \mathscr{B}_{p_{11,n}} & \mathscr{B}_{p_{21,n}} & \mathscr{B}_{p_{31,n}} \end{bmatrix},$$

i.e. probabilities vary not only in time n but also in the "space" position for fixed time. Notice that if $p_{9,n} = 1$ implies that central value is deterministically 0 which covers our model.

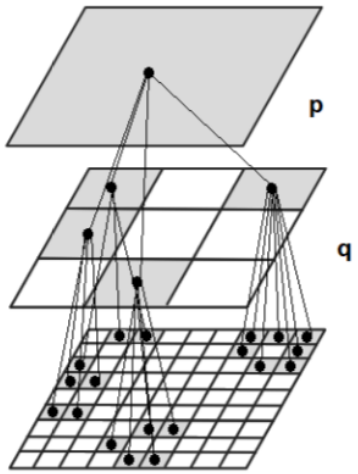
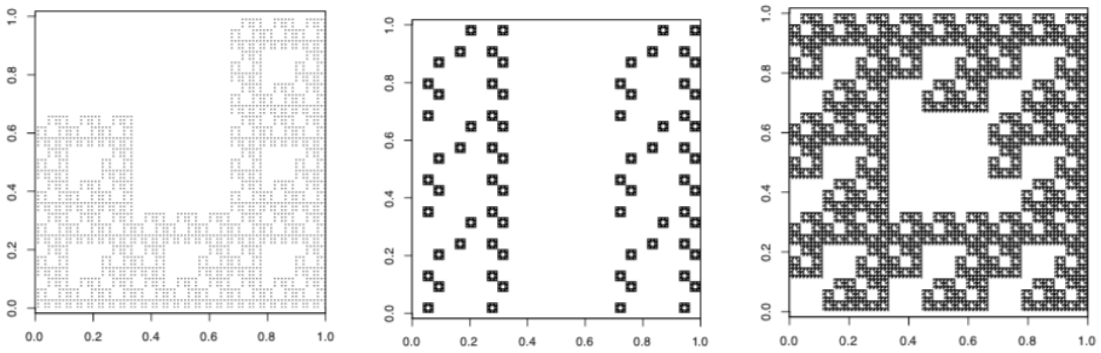


Fig. 1.4 Tree of $[pq]$ model.

1.4.2 IFS for random SC

The SC is a fractal subset of $\mathbb{R}^2$ which can be defined also as the fixed point of a family of contraction maps. If we take the original square to be a unit square with opposite corners at (0,0) and (1,1), then the IFS would be given by the following affine maps

$$F_i(\mathbf{x}) = \begin{bmatrix} 1/3 & 0 \\ 0 & 1/3 \end{bmatrix} \mathbf{x} + \frac{\mathbf{p}_i}{3}, \quad i = 1,\dots,8, \tag{1.5}$$



(a) $(0.9, 0.9, 0.9, 0.9, 0.1)$ (b) $(0.9, 0.9, 0.1, 0.9, 0.7)$ (c) $(0.9, 0.9, 0.9, 0.9, 0.7)$

Fig. 1.5 Random SC for $[p_1 p_2 p_3 p_4 p_5]$ model.

where

$$\mathbf{p}_i = ([0,0]^T, [0,1]^T, [0,2]^T, [1,0]^T, [1,2]^T, [2,0]^T, [2,1]^T, [2,2]^T).$$

This kind of random SC is formed by scaling an unit square by the factor $\frac{1}{3}$ and generated by following steps:

1) We pick random real number from interval $x_0 \in [0,1]$ and obtain a point $\mathbf{x}_0 := [x_0, x_0] \in [0,1]^2$.

2) Then we generate a value from the set $i \in \{1, \ldots, 8\}$ of values which are equally likely to be observed.

3) We obtain a value $\mathbf{x}_1 := F_i(\mathbf{x}_0)$.

4) We repeat this iteratively by generating of i in each step and using relation $\mathbf{x}_j := F_i(\mathbf{x}_{j-1}), \;\; j > 1$.

See Figure 1.6 where the results of specific number of iterations are plotted.

1.4.3 Random fractals and Kronecker product

Random fractal sets appear naturally from paths of stochastic processes or from random recursive constructions.

Here we deal with simple computation of matrix expressing the form of fractals, including random SC, in general. In (Dekking and Meester, 1990, Example 1.1) authors

(a) 1000 iterations (b) 3000 iterations (c) 10^4 iterations

Fig. 1.6 Random SC generated by IFS.

consider also a generalization of Mandelbrot's percolation process but this approach is different (depending on the set of indices generating zeros and ones). Let $\mathbf{A}$ be an $m \times n$ matrix and $\mathbf{B}$ be an $r \times s$ matrix. The Kronecker product of these matrices is defined as the $(mr) \times (ns)$ matrix

$$\mathbf{A} \otimes \mathbf{B} = \begin{bmatrix} a_{11}\mathbf{B} & \cdots & a_{1n}\mathbf{B} \\ \vdots & \ddots & \vdots \\ a_{m1}\mathbf{B} & \cdots & a_{mn}\mathbf{B} \end{bmatrix}.$$

It is associative, distributive but non-commutative. The multiple Kronecker product is defined in a recursive fashion as $\bigotimes_{i=1}^{r} A_i := \left(\bigotimes_{i=1}^{r-1} A_i\right) \otimes A_r$. Several known fractals can be represented by multiple Kronecker product, (Steeb, 2014b, chap. 9).

Example 1. *The fractal given by* $\bigotimes_{i=1}^{\infty} \mathbf{X}$ *can be thus represented by the "Matrix"* $\mathbf{X}$*. We list several known fractals.*

a) Cantor set - $\mathbf{X} = (1,0,1)^T$

b) Sierpiński triangle (Pascal's triangle mod 2) - $\mathbf{X} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$

c) Moiré-like pattern - $\mathbf{X} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$

For random variation we use the following notation $\mathbf{M}_r^{p_1,\ldots,p_r} := \bigotimes_{i=1}^{r} \mathbf{X}_{p_i}$, where r is natural number and $\mathbf{M}_r^{p_1,\ldots,p_\infty} := \bigotimes_{i=1}^{\infty} \mathbf{X}_{p_i}$. Here $\mathbf{X}_{p_i}$ means matrix involving values

generated by $\mathscr{B}_{\cdot,p_i}$. Evolving of random fractal can be defined followingly

$$\mathbf{Y}_0 := 1 \tag{1.6}$$

$$\mathbf{Y}_n := \mathbf{Y}_{n-1} \bigotimes \mathbf{X}_{p_n},\ n \geq 1. \tag{1.7}$$

Mutual independence implies directly $\mathbb{E}[\mathbf{Y}_n] = \mathbb{E}[\mathbf{Y}_{n-1}] \bigotimes \mathbb{E}[\mathbf{X}_{p_n}] = \bigotimes_{j=0}^{n} \mathbf{E}\left[\mathbf{X}_{p_j}\right]$, see Ghazal and Neudecker (2000). For a specific constant case let $\mathbf{M}_r^p := \mathbf{M}_r^{p,\dots,p}$, i.e. $\mathbf{M}_r^p = \bigotimes_{i=1}^{r} \mathbf{X}_{p_i},\ p_i = p$ for all $i \in \{1,\dots,r\}$.

Example 2. *For SC one can use matrix* $\mathbf{S} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 1 \end{bmatrix}$ *and for random SC* $\mathbf{S}_p = \begin{bmatrix} \mathscr{B}_{1,p} & \mathscr{B}_{2,p} & \mathscr{B}_{3,p} \\ \mathscr{B}_{4,p} & 0 & \mathscr{B}_{5,p} \\ \mathscr{B}_{6,p} & \mathscr{B}_{7,p} & \mathscr{B}_{8,p} \end{bmatrix}$.

Next lemma give us the mean value of retained elements since we know that the deterministic case the SC has exactly 2^{3n} elements in $n-$th iteration. One can directly obtain that for $X_{p_i}^j = \mathscr{B}_{j,p_i}$, $\tilde{p}_n = \prod_{j=1}^{n} p_j$ and $n \geq 1$ we have $\mathbf{Y}_n = \mathscr{B}_{\cdot,\tilde{p}_n}$ if in resulted deterministic matrix is value 1 and zero otherwise. Therefore for $k \in \mathbb{N}$, the $k-$th moment is $\mathbb{E}\left[(\mathbf{Y})_{\mathbf{n}}^{\mathbf{k}}\right] = \tilde{p}_n$ for values 1. Thus we have $N_n = 2^{3n}\tilde{p}_n$ retained elements in average, which reduces to $(8p)^n$ for case $p_j = p$. From that we see the balance threshold $p_b < \frac{1}{8}$ for which convergence to zero in average is obtained.

Using properties of Kronecker product and Frobenius norm one can show the following (Stehlík et al., 2018),

Theorem 7.

$$\|\mathbf{M}_r^{p_1,\dots,p_r}\|_F = \sqrt{\prod_{i=1}^{r} \mathrm{tr}\left(\mathbf{X}_{p_i}^* \mathbf{X}_{p_i}\right)}.$$

Under the assumption of independence we have $Z := \left\|\mathbf{M}_r^{p_1,\dots,p_r}\right\|_F^2 = \prod_{i=1}^{r} \sum_{j=1}^{8} \mathscr{B}_{j,p_i}^2 = \prod_{i=1}^{r} \mathrm{Bin}(8, \mathrm{p_i})$. Thus $\mathbb{E}[Z] = \prod_{i=1}^{r} 8p_i$, especially it is $(8p)^r$ for $p_i = p$. It means that for at least one small p_i the probability of Z value is close to zero. This probability $\mathscr{P}_{all}$ of event[1], that matrix $\mathbf{M}_r^p$ is zero matrix (empty) and probability $\mathscr{P}_{col}$ of event, that at least one column of matrix $\mathbf{M}_r^p$ is zero, has been estimated. Both but especially the latter is a specific percolation, since we can reach the bottom. The results are summarized in Table 1.1.

[1]Notice, that the probability $\mathscr{P}_{all}$, i.e. that Z is zero value is given by inclusion exclusion principle.

Table 1.1 Estimation of probabilities $\mathscr{P}_{all}, \mathscr{P}_{col}$ for given $1-p$ and r with 10^5 (10^4 for $r > 5$) repetitions.

$1-p \mid r$	1	2	3	4	5	6	7
0.9	0.43266 0.9854	0.67639 0.9996	0.81285 1	0.89705 1	0.93958 1	0.9666 1	0.9840 1
0.7	0.05842 0.7799	0.11173 0.9549	0.16310 0.9905	0.21126 0.9973	0.25628 0.9995	0.3022 1	0.3480 1
0.5	0.00345 0.4134	0.00733 0.6617	0.01187 0.8074	0.01558 0.8910	0.01944 0.9398	0.0221 0.9622	0.0340 0.9800
0.3	0.00001 0.1353	0.00010 0.2560	0.00021 0.3592	0.00025 0.4497	0.00026 0.5255	0.0006 0.5995	0.0010 0.6542

1.5 Fractal percolation

Fractal percolation provides a natural method of generating statistically self-similar fractals, with the same random process determining the form of the fractals at both small and large scales.

Random fractals can have complicated topological structure, for example they can be highly multiply connected. We consider topological aspects of a statistically self-similar construction known as "fractal percolation". In particular, this exhibits a "phase transition", where the topological structure changes dramatically when a continuously varying parameter increases through a critical value.

Best known is Mandelbrot percolation, based on repeated decomposition of squares into smaller subsquares from which a subset is selected at random. Let p be a number with $0 < p < 1$ and $E = E_0$ denote the unit square in $\mathbb{R}^2$. We divide the unit square E_0 into 9 squares of side $1/3$ in the obvious way. We select a subset of these squares to form E_1 in such a way that each square has independent probability p of being selected. Similarly, each square of E_1 is divided into 9 squares of side $1/9$, and each of these has independent probability p of being chosen to be a square of E_2. We continue in this way, so that E_k is a random collection of kth level squares of side 3^{-k}. This procedure, which depends on the parameter p, leads to the random *percolation set* and defines a random fractal $F_p = \bigcap_{k=0}^{\infty} E_k$. If $p > 3^{-2}$ then there is a positive probability of non-extinction, i.e. that $F_p \neq \varnothing$;, conditional on which $dim_H F_p = 2 + \log p / \log 3$ almost surely.

Proposition 1. *Givenp, let $t = q$ be the least positive solution of the equation*

$$t = (pt + 1 - p)^9.$$

Then F_p is empty with probability q. If $p \leqslant \frac{1}{9}$ then $q = 1$. If $\frac{1}{9} < p < 1$ then $0 < q < 1$ and, with probability $1 - q$, $dim_H F_p = dim_B F_p = \log 9p / \log 3$.

Proof. See demostration in Falconer (2004), pag. 252. □

We now discuss briefly the qualitative way in which the random set F_p changes as p increases from 0 to 1. We have already noted that F_p is almost surely empty if $0 < p \leqslant 1$. If $\frac{1}{9} < p < \frac{1}{3}$ we have, with probability 1, that either $F_p = \varnothing$ or $dim_H F_p = \log 9p / \log 3 < 1$, so by Proposition 2.5 F_p is totally disconnected. At the other extreme, if p is close to 1, it is plausible that such a high proportion of the squares are retained at each stage of the construction that F_p will connect the left and right sides of the square E_0; when this happens we say that percolation occurs between the sides. We show that this is the case at least if p is very close to 1; the ridiculous bound 0.999 obtained can certainly be reduced considerably.

Proposition 2. *Suppose that $0.999 < p < 1$. Then there is a positive probability (in fact bigger than 0.9999) that the random fractal F_p joins the left and right sides of E_0.*

Proof. See demostration in Falconer (2004), pag. 254. □

We have seen that if $0 < p < \frac{1}{3}$ then, with probability 1, F_p is empty or totally disconnected. On the other hand, if $p > 0.999$ then there is a high probability of percolation. The next theorem states that one or other of these situations pertains for each value of p.

1.6 Percolation threshold

The most known percolation model is to take some regular lattice and make it into a random network by randomly ”occupying” sites (vertices) or bonds (edges) with a statistically independent random variables. E.g. each of the lattice sites is either occupied (with probability p) or vacant (with probability $1 - p$). This is known as the site percolation problem. There is a well defined range of p for which the probability of percolation decreases rapidly from one to zero. It is centered on a critical value p_c called percolation threshold. Mandelbrot (1974) introduced a process in $[0,1]^2$ which he called ”canonical

curdling". It is the model from Remark 1 with $\mathscr{B}_{p_{ij,n}} = \mathscr{B}_p$. In paper Chayes et al. (1988) authors study the connectivity or "percolation" properties of such sets. They showed there is a probability $p_c \in (0,1)$ so that if $p < p_c$ then the set is "dustlike", whereas if $p \geq p_c$ opposing sides are connected with positive probability. We introduce now the exact definition of p_c. Let $B_n = \{x \in A_n : x$ can be connected to $[0,1] \times \{0\}$ and $[0,1] \times \{1\}$ by paths in $A_n\}$, $B_\infty = \bigcap_{n\in\mathbb{N}} B_n$ and let $\Omega = \{B_\infty \neq \emptyset\}$. When Ω occurs there is a up-to-down crossing of $[0,1]^2$. Finally we define $p_c = \inf_p\{\mathbb{P}(\Omega)\}$. Here we have to emphasize that for approximation of A_∞ we can not use this kind of definition. We just simply define p_c as the $\frac{1}{2}$ threshold of probability $\mathbb{P}(B_n = \emptyset)$.

The topological properties of Mandelbrot percolation have been studied extensively, see (Swishchuk and Islam, 2016; Falconer, 2004; Falconer and Grimmett, 1992) for surveys. In particular there is a critical probability p_c with $1/3 < pc < 1$ such that if $p > p_c$ then, conditional on non-extinction, F_p contains many connected components, so projections onto all lines automatically have positive Lebesgue measure. If $p \leqslant pc$ the percolation set F_p is totally disconnected (see Stehlík et al. (2018)).

Theorem 1. *There is a critical number p_c with $0.333 < p_c < 0.999$ such that if $0 < p < p_c$ then F_p is totally disconnected with probability 1, but if $p_c < p < 1$ then there is positive probability that F_p connects the left and right sides of E_0.*

Proof. Suppose p is such that there is a positive probability of F_p not being totally disconnected. Then there is positive probability of some two distinct points of F being joined by a path in F. This implies that there is a positive probability of the path passing through opposite sides of one of the squares in E_k for some k; by virtue of the statistical self-similarity of the construction, there is a positive probability of a path crossing E_0 from left to right. Clearly, if F_p has probability 1 of being totally disconnected, the same is true of $F_{p'}$ if $p' < p$. Thus the critical probability p_c is the supremum of those p such that F_p is totally disconnected, and has the properties stated. □

Experiment suggests that $0.7 < p_c < 0.8$.

1.6.1 Estimation of single parameter p

Here we consider $[p \dots p]$ model. The flow through the generated lattice (represented by the matrix) was used followed by a recursive depth first search (checking whether or not the flow makes it to the bottom of the grid). The graphs of a simulated data possess

Table 1.2 Estimation of percolation threshold p_c for given r .

r	1	2	3
est. p_c	0.34984	0.52455	0.61946

a sigmoidal shape which signals the presence of a critical value. Moreover we have categorical dependent variable represented by the outcomes pass/fail. All this indicates that in order to determine the threshold, we can fit a logistic model (log-odds)

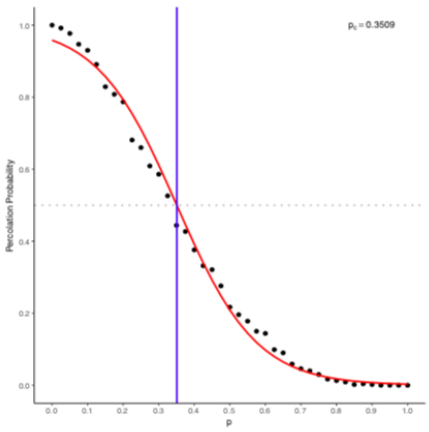
$$\ln\left(\frac{p}{1-p}\right) = \beta_0 + \beta_1 x.$$

It is fitted to simulated data and the threshold value for p (or $1-p$) is estimated such that $p = \frac{1}{2}$. But this implies $0 = \beta_0 + \beta_1 x$ and therefore estimation of p_c equals $-\frac{\beta_0}{\beta_1}$. Notice that similarly authors in Achter and Webb (2006) used logistic function to find threshold in order to examine the effects of mixed dispersal strategies on the spatial structure of a population considering a spatially explicit birth–death model. On the other hand, an experiment using the phenomenon of percolation has been conducted to demonstrate the implementation of neural functionality in Pignon et al. (1996), where the curve was found to be almost exactly described by the sigmoid form.

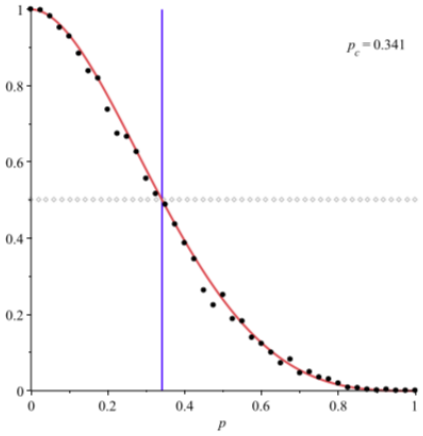
In Table 1.2 we give the results for $r = 1,2,3$, see also Figures 1.7 and 1.8, where the simulations of percolation probability for given p are shown. The logistic model seems to fit the curve very well. However for $r = 1$ (important generator case) we can explicitly find the analytic form of the model function. By going through all the possibilities we can directly find the polynomial

$$\mathscr{P}(p) = (1-p)^3\left[(1-p)^5 - 2(1-p)^4 + 2\right]. \tag{1.8}$$

The threshold value $\mathscr{P}(p) = \frac{1}{2}$ is given approximately as 0.341. See Figure 1.7 for excellent fitting and notice that logit model fits very closely with the value 0.34984.

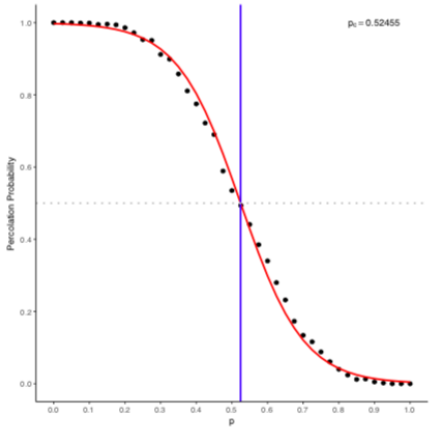


(a) Estimated logit model.

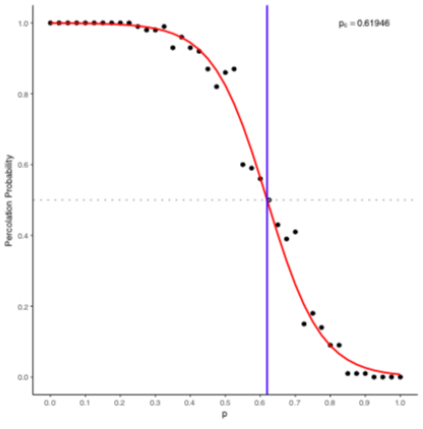


(b) Exact polynomial function.

Fig. 1.7 Percolation probability fitted to simulated data, $r = 1$.



(a) $r = 2$



(b) $r = 3$

Fig. 1.8 Percolation probability as a function of parameter p.

Chapter 2

Fractal and Texture Measure

2.1 Introduction

Calculating fractal dimensions is the primary objective in the study of fractals and can be a fairly complex task. Different methods or techniques have been propossed to estimate the FD of an image: (a) box-counting; (b) correlation; (c) spectral methods (Fourier spectrum); (d) sandbox; (e) diferential box counting; and (f) method wavelet; (g) others. The choosing of method for FD estimation depends on the nature of image. We are interested in grey level images because it is one of the most common representations of a given case study in medicine, science and engineering. For gray level images and its simplicity and automatic or semiautomatic computability the box-counting dimension and wavelet method are the most common methods based on a measure of area used for measurements in various application fields among other fractal techniques.

In this dissertation we pretend to understand the complexity of the structure of benign, malignant and normal breast cancer characterizing them through texture and fractal analysis. For this purpose we have made a complete review of the fractal methods, texture features values of GLCM (Grey Level Co-ocurrence Matrix), succolarity and lacunarity for applying to two-dimensional images of breast cancer. The most suitable methods for this task has been selected. The novelty of this paper is to analyze the efficency of the combining those techniques to differentiate between abnormal (cancerous) versus normal and betweeen normal, benign and malignant tissue on basis of mammogram images.

Scaling of tissue systems and its texture could be characterized with fractal, wavelet transform, lacunarity, succolarity, GLMC and fractal techniques. The fractal approach

assumes a hierarchical distribution of mass in space such that at any resolution the fractal structure is seen as the union of subjects similar to the whole. In this instance a single fractal dimension serve to characterize the mass distribution.

In this Chapter we introduce the methods used for calculating fractal and textural features measures. From the most simple box counting fractal method to multifractal gliding method for lacunarity. We extract statistical features from GLCM and we will classify images as normal, benign or malignant, analyzing the background of the image.

Fractals have been observed in nature and on computer screen and one of their excepcional characteristics is that they can be discribed by a noninteger dimension. Fractal analysis is now widely used in all areas of science with many applications in medicine. Fractal geometry can be useful to describe the pathological architecture of tumors, growth mechanism of blood vessels and can be applied for cancer detection. In mathematics, this concept has a long history and was introduced by Benoit Mandelbrot (B. Mandelbrot, 1983) based on his paper on self-similarity in which he discusssed the mathematical concept "*fractional dimensions*". Fractals are most often characterized by their fractional dimension (FD). The FD has been applied in texture analysis or description and segmentation (Hermann et al., 2015; Dobrescu et al., 2013), shape and classification (Nicolis et al., 2017, 2011; Ramírez-Cobo and Vidakovic, 2013; Stehlık et al., 2016), image and graphic analysis in many other fields (Abramovich, 2000; Derado et al., 2008; Davies and Hall, 1999). The breast tumors images are complex nature in its variation in textural behavior and to the noise existence. FD is a diagnosis tool helping to improve the sensitivity or the specificity of breast cancer. Fractal measures are the most frequently used in mammogram image analysis in application to mass detection (Dobrescu et al., 2013).

Since the fractal dimension describes how much space is filled, but does not indicate how the space is filled by the fractal structure, the lacunarity can solve this problem by making a distinction between different objects with the same fractal dimension. The lacunarity is a parameter that indicates the distribution of gap sizes throughout the object embedded in an image, it is able to identify different fractal structures that have the same fractal dimension. A textured image can be analysed with different methods, elements of fractal analysis (box-counting fractal dimension, lacunarity and succolarity) proving to be a good choice for texture analysis and classification (Popescu et al., 2017). Succolarity prove to be a good feature for texture classification problems (Cojocaru et al., 2013; De Melo and Conci, 2008, 2013; Popescu et al., 2017). Succolarity is the fractal property that presents the percolation degree of a fluid flowing around and through an object in an image on a

given direction. All three elements of fractal analysis were first theoretically described in (B. Mandelbrot, 1983), but only fractal dimension had been calculated. Popescu et al. (2017) and De Melo and Conci (2013) present an efficient method to compute succolarity and Cojocaru et al. (2013) takes into account this element of fractal analysis as an interesting texture descriptor used for texture classification.

2.2 Hausdorff dimension

The first definition of fractal dimension is the Hausdorff dimension. Hausdorff dimension provides a rigorous mathematical definition of dimension (B. Mandelbrot, 1983). As mentioned in the definition (1), the Hausdorff dimension or Hausdorff-Besicovitch dimension is a metric generalization of the dimension concept of a topological space, which allows defining a fractional dimension for a fractal object. It is the oldest dimension, among the variety of fractal dimensions in use and presents the advantage of being defined for any set. It is also mathematically convenient, as it is based on measures, which are relatively easy to manipulate. However, the comprehension of this dimension is essential to understand the mathematics of fractals.

In an intuitive way, Hausdorff dimension can be introduced through the exponent describing the variation of the size of an object with the scale used to measure it (B. Mandelbrot, 1983; Theiler, 1990),

$$\text{size} \sim \text{scale}^{\text{dimension}} \tag{2.1}$$

For a segment, both its size and scale are given by its length and the dimension is one. A fractal object in the plane, like a coastline, will have dimension larger than one (and smaller than two) as a consequence of the space-filling properties of the graph and its infinite length.

Returning to definition (4) and equation (1.1) it is clear that for any given set $F \in \mathbb{R}^n$ and $\delta < 1$, $\mathscr{H}^s_\delta(F)$ is non-increasing with s, so by (1.2) $\mathscr{H}^s(F)$ is also non-increasing. In fact, rather more is true: if $t > s$ and U_i is a $\delta-$cover of F we have

$$\sum_i |U_i|^t \leqslant \sum_i |U_i|^{t-s} |U_i|^s \leqslant \delta^{t-s} \sum_i |U_i|^s \tag{2.2}$$

so, taking infima over all $\delta-$covers,

$$\mathscr{H}^s_\delta(F) \leqslant \delta^{t-s}\mathscr{H}^s_\delta(F)$$

Letting $\delta \to 0$, we see that if$\mathscr{H}^s(F) < \infty$, then $\mathscr{H}^t(F) = 0$ for $t > s$.

Before defining the dimension of Hausdosrff or fractal dimension, let's observe the following:

For any given set $F \in \mathbb{R}^n$, $\delta > 0$ and $0 \leq r < t < \infty$, we have that

$$\begin{aligned}
\mathscr{H}^t_\delta &= \inf\left\{\sum_{i=1}^{\infty}(diam(S_i))^t : F \subseteq \bigcup_{i\in I} S_i \ \ and \ \ diam(S_i) \leq \delta\right\} \\
&= \inf\left\{\sum_{i=1}^{\infty}(diam(S_i))^s (diam(S_i))^{t-s} : F \subseteq \bigcup_{i\in I} S_i \ \ and \ \ diam(S_i) \leq \delta\right\} \\
&\leq \delta^{t-s}\inf\left\{\sum_{i=1}^{\infty}(diam(S_i))^s : F \subseteq \bigcup_{i\in I} S_i \ \ and \ \ diam(S_i) \leq \delta\right\} \\
&= \delta^{t-s}\mathscr{H}^r_\delta(F)
\end{aligned}$$

i.e, $\mathscr{H}^t_\delta(F) = \delta^{t-s}\mathscr{H}^r_\delta(F)$ taking limits when $\delta \to 0$ we have that: $\mathscr{H}^s(F) < \infty$ it implies $\mathscr{H}^t(F) = \infty$ for $t < s$ and $\mathscr{H}^t(F) = 0$ for $t > s$, therefore there is a single value s associated with the set F such that

$$\begin{cases}
\mathscr{H}^t\ (F) = \infty \quad if\ \ t < s \\
\mathscr{H}^t(F) = 0 \quad if\ \ t > s
\end{cases}$$

Taking into account the previous observation we can define the dimension of Hausdorff.

Definition 15. *The critical value s is called the Hausdorff dimension or fractal of F, and written* $\dim_H F$*; it is defined for any set* $F \subset \mathbb{R}^n$*. Formally in (Falconer, 2004, 1997)*

$$\dim_H F = \inf\{s \geq 0 : \mathscr{H}^s(F) = 0\} = \sup\{s : \mathscr{H}^s(F) = \infty\} \tag{2.3}$$

(taking the supremum of the empty set to be 0), so that

$$\mathscr{H}^s(F) = \begin{cases}
\infty \quad if\ \ 0 \leq s < \dim_H F \\
0 \quad if\ \ s > \dim_H F
\end{cases}$$

If $s = \dim_H F$, then $\mathscr{H}^s(F)$ may be zero or infinite, or may satisfy

$$0 < \mathscr{H}^s(F) < \infty$$

Hausdorff dimension satisfies the following properties, which will allow us to calculate the fractal dimension of fractal objects, for example, the SC.

1. **Monotonicity**. If $E \subset F$ then $\dim_H E \leqslant \dim_H F$. This is immediate from the measure property that $\mathscr{H}^s(E) \leqslant \mathscr{H}^s(F)$ for each s.

2. **Countable stability**. If $F_1, F_2, \ldots$ is a (countable) sequence of sets then $\dim_H \bigcup_{i=1}^{\infty} F_i = \sup_{1 \leqslant i \leqslant \infty} \{\dim_H F_i\}$. Certainly, $\dim_H F_i \geqslant \dim_H F_j$ for each j from the monotonicity property. On the other hand, if $s > \dim_H F_i$ for all i, then $\mathscr{H}^s(F_i) = 0$, so that $\mathscr{H}^s\left(\bigcup_{i=1}^{\infty} F_i\right) = 0$, giving the opposite inequality.

3. **Countable sets**. If F is countable then $\dim_H F = 0$. For if F_i is a single point, $\mathscr{H}^0(F_i) = 1$ and $\dim_H F_i = 0$, so by countable stability $\dim_H \bigcup_{i=1}^{\infty} F_i = 0$.

4. **Open sets**. If $F \subset \mathbb{R}^n$ is open, then $\dim_H F = n$. Since F contains a ball of positive $n-$dimensional volume, $\dim_H F \geqslant n$, but since F is contained in countably many balls, $\dim_H F \leqslant n$ using countable stability and monotonicity.

5. **Smooth sets**. If F is a smooth (i.e. continuously differentiable) $m-$dimensional submanifold (i.e. $m-$dimensional surface) of $\mathbb{R}^n$ then $\dim_H F = m$. In particular smooth curves have dimension 1 and smooth surfaces have dimension 2. Essentially, this may be deduced from the relationship between Hausdorff and Lebesgue measures.

6. **Hölder condition**. Let $F \subset \mathbb{R}^n$ and suppose that $f : F \to m$, then $|f(x) - f(y)| \leqslant c\,|x-y|^{\alpha}$ with $(x, y \in F)$. Then $\dim_H f(F) \leqslant (1/\alpha)\dim_H F$.

7. **Lipschitz transformation**. If f is a Lipschitz transformation then $\dim_H f(F) \leqslant \dim_H F$.

Suppose now that $s = \dim_H(F)$ which implies that $\mathscr{H}^S(F) < 1$, that there is a similarity f of ratio r and also that we can calculate the proportion p of the measure of F with respect to the measure of $f(F)$; that is to say:

$$\begin{cases} \mathscr{H}^s(f(F)) = r^s \mathscr{H}^s(F) \\ \mathscr{H}^s(f(F)) = p\mathscr{H}^s(F) \end{cases}$$

then $r^s = p$, thus

$$s = \frac{\log p}{\log r} \tag{2.4}$$

Let us use this last equation to calculate the dimension of Hausdorff of the SC.

Example 3 (Hausdorff dimension of the SC)**.** *Let F be the SC set (see Figure 1.3).Then $dim_H F = \frac{\log 8}{\log 3} = 1,89\ldots$ with $\frac{1}{8} \leqslant \mathscr{H}^s(F) \leqslant 1$ where $s = \frac{\log 8}{\log 3}$.*

The equation 2.4 can be used to calculate the dimension of all fractals that we know the rules of construction and are composed with parts of same proportion. Let's see a example, now on 2D: Sierpinski Carpet (B. Mandelbrot, 1983) which construction scheme is represented in figure 1.3.

Let's remember the SC definition (14), to divide the square in nine equal squares we use a scale factor of $r = 1/3$. The number of pieces to reconstruct is eight (because we cut out only the middle one), so $N = 8$. Then, the dimension of the Sierpinski Carpet (figure 1.3) is:

$$s = \frac{\log p}{\log r} = \frac{\log 8}{\log 3} \cong 1,89$$

A fractal dimension for a 3D figure, like a cube, can also be computed using this formula as we see on the next example for the fractal Menger's Sponge (B. Mandelbrot, 1983) , also knew as Sierpinski Sponge, built by Karl Merger in 1926, is a generalization of the three-dimensional Sierpinski Carpet, that is, instead of an initial square, there is an initial cube and instead of removing squares, cubes are removed. Its construction start with a cube and at each stage divide each remaining cube into twenty seven equal cubes and cut out the central one and all the other six that share faces with it and. This procedure is represented in figure 2.1.

Fig. 2.1 Construction of the Menger's sponge.

This example presents an idea similar to the Sierpinski Carpet dividing the cubes with a scale factor of $r = 1/3$. But now, the number of pieces to reconstruct is twenty (because we cut out seven pieces: the middle one and other six), so $N = 20$. Then, the dimension of the Menger's Sponge (figure 2.1) is:

$$s = \frac{\log p}{\log r} = \frac{\log 20}{\log 3} \cong 2,73$$

2.3 Similarity and fractal dimension

One of the most intuitive notions of dimension is associated with scale and self- similarity. A 1D object (of dimension 1), a line for example, if divided in N parts, each part will be equal to the previous one scaled by the factor $\delta = 1/N$, and $N \times \delta^1$ will reconstruct the object.

A 2D object (of dimension 2), a square, for example, could be divided in N equal parts, each will be equal to the previous one scaled by $\delta = \sqrt{\frac{1}{N}} = \left(\frac{1}{N}\right)^{\frac{1}{2}}$ and $N \times \delta^2$ reconstructs the object.

An example in 3D, a cube for example, could be divided in N equal parts, each will be equal to the previous one scaled by $\delta = \sqrt[3]{\frac{1}{N}} = \left(\frac{1}{N}\right)^{\frac{1}{3}}$ and $N \times \delta^3$ reconstructs the object.

The powers of δ in each case are the dimensions D of the objects. On fractals, D is fractional. This definition of dimension, denoted by D_S, that is D with the subscript 'S' , is named similarity dimension.

So, the dimension D_S should be Falconer (1997, 2004):

$$N \times (\delta)^{D_S} = 1$$

$$N = \frac{1}{\delta^{D_s}} = \left(\frac{1}{\delta}\right)^{D_s} \Longrightarrow \log N = D_s \times \log\left(\frac{1}{\delta}\right)$$

$$D_s = \frac{\log N}{\log\left(\frac{1}{\delta}\right)}$$

where, N is the number of parts to reconstruct the original figure, and δ is the ratio between consecutive steps.

Although this way to measure the Fractal Dimension (FD) is extremely simple, it only could be applied when we know the rules of construction of the object (as observed earlier) and all parts have the same relations. This occurs only on the so called deterministic fractals. When more complex objects are under consideration as natural elements like: clouds, trees, neurons, dusts, tissues, wave frequency, color radiation, surfaces of the sea, etc; then we have lots of other ways to measure the fractal dimension experimentally as we will see next.

Definition 16. *The FD quantifies the density of fractals or any images on its metric space. The fractal dimensions are then an objective way to compare one fractal (or figure) to another.*

However a list of characteristic properties of a FD can be given (B. Mandelbrot, 1983; Falconer, 1986, 1997, 2004). A fractal dimension $FD(F) \in \mathbb{R}^n$ of a set F has at least the following properties:

1. **motion invariance**: $FD(f(F)) = FD(F)$, if f is a translation or a rotation
2. **scale** invariance: $FD(f(F)) = FD(F)$, if f is a similarity or an affine transformation
3. **monotony**: if $E \subseteq G$, then $FD(E) \leqslant FD(G)$
4. **maximum property**: $FD(E \cup G) = \max\{FD(E), FD(G)\}$
5. **extension of topological dimension**: If M is a smooth $m-$dimensional manifold (curve, surface, etc.), then $FD(M) = m$.

Fractal dimension characterizes the irregularity or complexity of the whole set under investigation and is governed by the most complex area of the set because of property 4.

The relation (2.4) leads to several important methods of estimating the FD for a given set F as we will present below.

2.4 Dimension Estimates

When applied in image analysis, fractal geometry is often brought to the evaluation of the FD. There are many definitions or approaches of calculating *Fractal Dimension (FD)* and each method has its own theoretic basis. This fact often leads to obtain different dimensions by different methods for the same feature. These differences appear because the Hausdorff–Besicovitch dimension (Eq. 2.4) is not computable in this form in most cases. Thus, the methods approximate it using different algorithms to estimate the parameter N. In this section we will see some of them, the most widespread literature, including: the box counting dimension (D_B); the information dimension (D_I)); correlation dimension (D_C); differential box-counting (DBC) and fractional Brownian motion (fBm). Although, the applied algorithms differ, they obey to the same basis summarized by the three steps:

- Measure the quantities of the object using various step sizes.
- Plot log (measured quantities) versus log (step sizes) and fit a least-squares regression line through the data points.
- Estimate FD as the slope of the regression line.

The first methods were the box-counting methods (Lopes and Betrouni, 2009). In these methods the signals are represented on a finite scale grid and the grid effects interplayed with the computing fractal dimension. For this reason, other methods have been defined to remove the grid effects such as the fBm methods.

2.4.1 Box-counting dimension

For most rigorously self-similar subsets encountered, Hausdorff and box-counting dimensions are in fact the same thing Salat et al. (2017). Box-counting or box dimension is one of the most widely used dimensions and conceptually the simplest dimension in use. Its popularity is largely due to its relative ease of mathematical calculation and empirical estimation. Although, this method was defined by Russell et al. (1980), the definition goes back at least to the 1930s and it has been variously termed Kolmogorov entropy, entropy dimension, capacity dimension (a term best avoided in view of potential theoretic associations), metric dimension, logarithmic density and information dimension. We shall always refer to box or box-counting dimension to avoid confusion.

Fractals reveal how irregular objects can be interpreted and described according to morphological characteristics. The irregular behaviors of these complex structures are difficult or impossible to be quantified by standard modeling techniques or by traditional Euclidean geometry. Traditional box-counting method "covers" the object with rectangular coordinate grid (squares for the plane and cubes in space) and counts the number of boxes. As each set of boxes is characterized by the square side r, the corresponding number of squares N of nonempty grid boxes necessary to cover the pattern is presented as a function of r. Thus, assuming that any fractal object is intrinsically self-similar, a simplified FD (i.e., box dimension) is determined as the slope of the $\log - \log$ relationship between $N(r)$ and r.

$$D_B = -\frac{\log(N)}{\log(r)} \tag{2.5}$$

where D_B is the box dimension, N is number of rules with linear length r used to cover the object.

Strictly mathematically "speaking", see Falconer (1997, 2004), for F a non-empty bounded subset of $\mathbb{R}^n$ let $N_r(F)$ be the smallest number os sets of diameter r that can cover F. (Recall definition 3 that the *diameter* of a set U is defined as $|U| = \sup\{|x-y| : x,y \in U\}$ that is the greatest distance apart of any pair of points in U.) The *lower* and *upper box-counting (or box) dimensions* of a subset $F \subset R^n$ are, respectively, defined by

$$\overline{\dim}_B(F) = \underline{\lim}_{r\to 0} \frac{\log N_r(F)}{-\log r}, \quad \underline{\dim}_B(F) = \overline{\lim}_{r\to 0} \frac{\log N_r(F)}{-\log r} \tag{2.6}$$

or $N(r) \sim (1/r)^{\dim_B}$. Of course, $\overline{\dim}_B(F) \geq \underline{\dim}_B(F)$ and if these values are equal, then the common value is referred to as the *box-counting dimension* or *box dimension* of F and is denoted by

$$\dim_B(F) = \lim_{r\to 0} \frac{\log N_r(F)}{-\log r} = \lim_{r\to 0} \frac{\log N_r(F)}{\log \frac{1}{r}} \tag{2.7}$$

where $N_r(F)$ can be the smallest number of cubes of side r (naturally, in 3D) or squares (in 2D) that covers F or the largest number of disjoint cubes or squares of side r with centers in F. Here, and trhoughout the this work, we assume that $r > 0$ is sufficiently small to ensure that $-\log r$ and similar quantities are strictly positive.

In the context of equation (2.6), box dimension is sometimes referred to as *Minkowski dimension* or *Minkowski–Bouligand dimension*. It is important to understand the relationship between box-counting dimension and Hausdorff dimension. If F can be covered by $N_r(F)$ sets of diameter r, then, from definition (4),

$$\mathscr{H}^s_r(F) \leqslant N_r(F) r^s$$

.

If $1 < \mathscr{H}^s(F) = \lim_{r\to 0} \mathscr{H}^s_r(F)$ then $\log N_r(F) + s \log r > 0$ if r is sufficiently small. Thus $s \leq \lim_{r\to 0} \frac{\log N_r(F)}{-\log r}$ so

$$\overline{\dim}_B(F) \geq \underline{\dim}_B(F) \geq \dim_H \tag{2.8}$$

for every $F \subset \mathbb{R}^n$.

Roughly speaking, 2.7 says $N_r(F) \simeq cr^{-s}$ for small r, where $s = dim_B F$, or, more precisely, that

$$N_r(F) r^s \to \infty \quad if \;\; s < \dim_B F$$

and

$$N_r(F) r^s \to 0 \quad if \;\; s > \dim_B F$$

There are several equivalent definitions of box dimension that are sometimes more convenient to use. Since box dimensions are determined by coverings by sets of equal size they tend to be easier to calculate than Hausdorff dimensions. Just as with Hausdorff dimension, calculations of box dimension usually involve finding a lower bound and an upper bound separately, each bound depending on a geometric observation followed by an algebraic estimate.

Example 4 (Box-counting dimension of the Sierpinski carpet)**.** *The next procedure will allow us to find the box-counting dimension of the Sierpinski carpet (see (Falconer, 2004; Hermann et al., 2015)).*

- *We will use* $2-$*dimensional boxes of size* $r = 3^{-n}$.
- *From the construction of the Sierpinski carpet, we can cover it using* 8^n *boxes of size* 3^{-n} *(since in fact the nth stage of the construction consists of exactly* 8^n *solid boxes each of side length* 3^{-n}*). Furthermore, it is not possible to use fewer boxes of this size, so* $N(3^{-n}) = 8^n$.

- *Thus the limit is* $\lim_{r\to\infty} \frac{\log(8^n)}{\log(1/3^{-n})} = \lim_{r\to\infty} \frac{n\log(8)}{n\log(3)} = \frac{\log 8}{\log 3} = \log_3 8$
- *So the box-counting dimension of the Sierpinski carpet is* $\log_3 8$.

2.4.2 Information dimension

Another way to estimate the dimension is called the information dimension (D_I) (Theiler, 1990; Addison, 1997). In Pitsianis et al. (1989) was introduced the concept of the information dimension D_I for transport processes on lattices. The idea is similar to the box counting dimension. The set is also covered with squares of side length r (squares for the plane and cubes in space). But this measure takes into account how much of the set is contained within each square instead of simply counting the squares (Falconer, 1997, 2004; Addison, 1997).

This alternative approach determines how much of the set (or attractor) is contained within each square $H(r)$. This measure seeks to account for differences in the distribution density of points covering the set, and is then given by

$$\dim_I(F) = \lim_{r\to 0} \frac{H_r(F)}{-\log r} \tag{2.9}$$

H is the Shannon's Entropy of the fractal. If we partition the fractal in boxes of size r we need

$$H = -\sum_i P_i \log P_i \tag{2.10}$$

where P_i is the probability measure (the size) of box i with side length r. For the special case of an set with an even distribution of points, an identical probability, $P_i = 1/N$, is associated with every box. Hence, $H_r = \log N_r(F)$, and the equation (2.9) reduces to the box counting dimension equation (2.7). Thus, D_B simply counts all cubes containing parts of the set, while D_I measure how much of the set is within each cube.

2.4.3 Correlation dimension

An alternative to estimate the dimension that requires less computation time is the correlation dimension (D_C). To define D_C is necessary that we first define the correlation sum C_r, as follows ((Addison, 1997; Liebovitch and Toth, 1989):

$$C_r = \frac{1}{N(N-1)} \sum_{i=1}^{N} \sum_{j=1; j\neq i}^{N} \theta\left(r - \left|X_i - X_j\right|\right) \tag{2.11}$$

where θ is the Heaviside function, r is the radius of an $n-$dimensional sphere centered on each sampled point on the trajectory of the set, X_i, $i = 1,2,3,\ldots,N$. X_i denotes the multidimensional vector that is the ith phase space coordinate of the set, i.e. for a set in 3D phase space with coordinates x, y and z, $X_i = (x_i, y_i, z_i)$. The Heaviside function counts the number of points of the set within the sphere centered on one point (that is not counted).

For a system of n first-order ordinary differential equations, the $2\,n-$dimensional space consisting of the possible values of $(x_1, x_1, x_2, x_2, \ldots, x_n, x_n)$ is known as its phase space.

The Heaviside function is equal to the unity if the value inside the brackets is positive and zero if the value is negative. X_i are the points on the reference trajectory and X_j are other points on the set in the vicinity of Xi. $\left|X_i - X_j\right|$ is the separation distance between the two points. $\theta\left(r - \left|X_i - X_j\right|\right)$ return a value of unity if the distance between the two points is less than the sphere radius; otherwise a zero value is returned. The calculation of the correlation sum of equation (2.11) involves following the reference trajectory, stopping at each discrete point on this trajectory and counting the number of other set points within a sphere of radius r. The cumulative sum of all the counted points is then divided by $N(N-1)$ to give the correlation sum, C_r. The maximum value of C_r is unity.

The correlation sum scales with the sphere radius according to a power law of the form:

$$C_r \propto r^{D_C}$$

where the exponent (D_C) is the correlation dimension, and is then given by

$$\dim_C(F) = \lim_{r\to 0} \frac{C_r(F)}{-\log r} \tag{2.12}$$

Examining the set in the method described for many different sphere radii, D_C is obtained from the slope of the scaling region of a $\log(r) \times \log(C_r)$ plot. In general, good estimates of D_C require large number of data points, N; however, the computational time required to calculate C_r increases with N^2.

The *correlation dimension*, *information dimension* and *box counting dimension* are related. In fact, they are part of a generalized collection of dimensions, known as Renyi dimensions. Theiler (1990) has defined a generalized information measure:

$$S_q(r) = \frac{1}{q-1} \log \sum_i P_i^q \tag{2.13}$$

which reduces to Shannon's formula (2.10) in the limit $q \to 1$, where $S(r)$ is the information (in bits) needed to specify a point on the set to an accuracy r. The generalized information dimension associated with the Renyi entropy is just the generalized dimension.

Based on the work of Rényi (1955) we can calculate a (generalized) fractal dimension by the formula:

$$D_q = \frac{1}{1-q} \lim_{r \to 0} \frac{\log \sum_{i=1}^{N(r)} P_i^q(r)}{log\left(\frac{1}{r}\right)} \tag{2.14}$$

where D_q, the generalized fractal dimension; q, dimension index ($q = 0$ for the generic fractal dimension (D_B)); P_i, the mass located within the box.

It can be shown that D_C forms a lower bound to D_I which in turn forms a lower bound to D_B,

$$D_C \leq D_I \leq D_B$$

In addition, D_C and D_I asymptotically approach the value of D_B as the attracting set becomes more uniformly distributed in phase space. If the points on the attractor are uniformly distributed, then $D_C = D_I = D_B$. In many practical cases D_C gives a very close estimate of D_B.

2.4.4 Differential Box-counting

Other method of measuring FD is the differential box-counting (DBC) (Liebovitch and Toth, 1989) proposed by Sarkar and Chaudhuri (1992), which, instead of counting the boxes like the previous box counting method, divide the image of size $M \times M$ into grids of $s \times s$ pixels. They shows that the DBC method, described in their paper is an adaptation of the box-counting method, which was proposed to solve some of limitations of the first method. DBC is the best in terms of dynamic range of FD values and efficiency, and also has good accuracy and generality.

It has the advantage to work on grey-scale images and thus the binarization step is avoided. The signal is partitioned into boxes of various size r and $N(r)$ is computed like

the difference between the minimum and the maximum grey levels in the (i, j)th box. This step is repeated for all boxes and the FD is estimated as in Eq. (2.7).

Below is a description of the algorithm step by step.

This method (Jin et al., 1995) considers an image of size $M \times M$ which has been scaled down to size $s \times s$, where s ranges between 1 to $M/2$, s is an integer. Then, $r = s/M$. The image is viewed as a 3-D space with x, y denoting 2-D position and the third coordinate z denoting gray-level. The x, y space is partitioned into grids of size $s \times s$. On each grid there is a column of boxes of size $s \times s \times h$ where h is the height of a single box. If the total number of gray-levels is G, then $G/h = M/s$. Let the minimum and maximum gray-levels of the image in the $(i, j)th$ grid fall in box number $k$and l, respectively. Then $n_r(i, j) = l - k + 1$ is the contribution of N_r in the $(i, j)th$ grid. Due to the differential nature in computing n_r this method is called differential box counting method.Taking contributions from all grids, we have

$$N_r = \sum_{i,j} n_r(i, j)$$

where N_r is counted for different values of s i.e. different values of r. Then using

$$\dim_{DBC}(F) = \lim_{r \to 0} \frac{N_r(F)}{-\log r} \tag{2.15}$$

the FD can be estimated from the least squares linear fit of $\log(N_r)$ against $\log(1/r)$. The slope of the best fitting curve will give the fractal dimension.

2.5 Wavelet Analysis

The use of wavelet analysis combined with fractional Brownian motion (fBm) analysis may provide an interesting approach to compute key values for fBm processes, such as the fractal dimension FD (Parra et al., 2003). In this thesis we propose 2D model to calculate the Hurst Coefficient H (and hence FD) for two-dimensional signals. This is based on their spectral characteristics. A formal extension of these model to 2D processes is developed (see Appendice B) and implemented. The proposed algorithm is tested using real data.

To define an wavelet based estimation of the FD we need the definition of a Fractional Brownian Motion (fBm) and some properties.

2.5.1 Fractional Brownian Motion

Fractional Brownian Motion (fBm) has been successfully exploited to model an important number of physical phenomena and non-stationary processes such as medical images. These mathematical models closely describe essential properties of natural phenomena, such as self-similarity,scale invariance and fractal dimension (FD).

Brownian motion is closely linked to the normal distribution. Recall that a random variable X is normally distributed with mean μ and variance σ^2 if

$$\mathbb{P}\{X > x\} = \frac{1}{\sqrt{2\pi\sigma^2}} \int_x^{\infty} e^{-} \frac{(u-\mu)^2}{2\sigma^2} du \quad for\ all\ x \in \mathbb{R}$$

The fractional Brownian motion (fBm) is a continuous Gaussian process X_t, depending on a parameter $H \in (0,1)$ called the Hurst index. It has expected value 0 and covariances given by

$$\langle X_t X_s \rangle = R_H(t,s) = \frac{1}{2}\left(|t|^{2H} + |s|^{2H} - |t-s|^{2H}\right) \tag{2.16}$$

The fBm is scale-invariant, namely $\left(a^{-H} X_{at}\right)_{t\geq 0}$ has the same law as $(X_t)_{t\geq 0}$ for all $a > 0$.

This process was introduced by Kolmogorov (Kolmogorov, 1940) and studied by Mandelbrot and Van Ness in Mandelbrot and Van Ness (1968), where a stochastic integral representation in terms of a standard Brownian motion was established. The parameter H is called *Hurst index* from the statistical analysis, developed by the climatologist Hurst (Hurst, 1951), of the yearly water run-offs of Nile river.

Definition 17. *A real-valued stochastic process $\{B(t) : t \geq 0\}$ is called a **(linear) Brownian motion** with start in $x \in \mathbb{R}$ if the following holds:*

- $B(0) = x$,
- *the process has independent increments, i.e. for all times $0 \leq t_1 \leq t_2 \leq \ldots \leq t_n$ the increments $B(t_n) - B(t_{n-1})$, $B(t_{n-1}) - B(t_{n-2}), \ldots, B(t_2) - B(t_1)$ are independent random variables,*
- *for all $t \geq 0$ and $h > 0$, the increments $B(t+h) - B(t)$ are normally distributed with expectation zero and variance h,*

- *almost surely, the function $t \to B(t)$ is continuous.*

*We say that $\{B(t) : t \geq 0\}$ is a **standard Brownian motion** if $x = 0$.*

Definition 18. *A stochastic process $\{B(t) : t \geq 0\}$ in $\mathbb{R}^n$ is called an $n-$**dimensional Brownian motion** started at $(x_1, \ldots, x_n)$ if*

$$B(t) = (B_1(t),\ B_2(t), \ldots, B_n(t))$$

where $B_1, \ldots, B_n$ are independent linear (scalar) fractional Brownian motions started at $x_1, \ldots, x_n$ respectively. An $n-$dimensional Brownian motion started at the origin is called standard.

Theorem 1 (Scaling invariance)**.** *Suppose $B(t) : t \geq 0$ is a standard Brownian motion and let $\alpha > 0$. Then the process $X(t) : t \geq 0$ defined by $X(t) = \frac{1}{\alpha} B(\alpha^2 t)$ is also a standard Broownian motion (see (Mandelbrot and Van Ness, 1968)).*

Proof. The properties of continuity and independence of increments remain unchanged. Additionally, $X(t) - X(s) = \frac{1}{\alpha}(B(\alpha^2 t) - B(\alpha^2 s))$ is normally distributed with expectation zero and variance $\frac{1}{\alpha^2}(\alpha^2 t - \alpha^2 s) = t - s$. □

The fractal properties of fBm have been studied extensively (see (Mandelbrot and Van Ness, 1968; Adler, 1981; Nualart, 2006; Xiao, 2013)). Its theoretical interest lies, in particular, in the fact that it is (up to normalization) the only Gaussian process satisfying the following two properties:

1. *Self-similarity:* For any constant $a > 0$, the processes $a^{-H} B^H_{at}, t \geq 0$ and $B^H_t, t \geq 0$ have the same probability distribution. This property is an immediate consequence of the fact that the covariance function 2.16 is homogeneous of order $2H$, and it can be considered as a "fractal property" in probability.

2. *Stationary increments:* From 2.16 it follows that the increment of the process in an interval $[s,t]$ has a normal distribution with zero mean and variance

$$\mathbb{E}\left(\left(B^H_t - B^H_s\right)^2\right) = \mathbb{E}\left[(B_{t+H} - B_{s+H})(B_{u+H} - B_{v+H})\right] = |t-s|^{2H} \tag{2.17}$$

for every $H, s, t, u, v \in \mathbb{R}$.

Hence, for any integer $k \geq 1$ we have

$$\mathbb{E}\left(\left(B_t^H - B_s^H\right)^2\right) = \frac{(2k)!}{k!2^k}\left|t-s\right|^{2k} \tag{2.18}$$

Choosing k such that $2Hk > 1$, Kolmogorov's continuity criterion and 2.18 imply that there exists a version of the fBm with continuous trajectories. Moreover, using Garsia-Rodemich-Rumsey lemma (Garsia et al., 1970), we can deduce the following modulus of continuity for the trajectories of fBm: For all $\varepsilon > 0$ and $T > 0$, there exists a nonnegative random variable $G_{\varepsilon,T}$such that $\mathbb{E}\left(\left|G_{\varepsilon,T}\right|^p\right) < \infty$ for all $p \geq 1$, and, almost surely,

$$\left|B_t^H - B_s^H\right| \leq G_{\varepsilon,T}\left|t-s\right|^{H-\varepsilon},$$

for all $s,t \in [0,T]$. In other words, the parameter H controls the regularity of the trajectories, which are Hölder continuous of order $H-\varepsilon$, for any $\varepsilon > 0$. For $H = 1/2$, the covariance 2.16 can be written as $R_{1/2}(t,s) = min(s,t)$, and the process $B^{1/2}$ is an ordinary Brownian motion. In this case the increments of the process in disjoint intervals are independent. However, for $H \neq 1/2$, the increments are not independent.

Set $X_n = B_n^H - B_{n-1}^H, n \geq 1$. Then $X_n, n \geq 1$ is a Gaussian stationary sequence with unit variance and covariance function

$$\begin{aligned}\rho_H(n) &= \frac{1}{2}\left((n+1)^{2H} + (n-1)^{2H} - 2n^{2H}\right)\\ &\approx H(2H-1)\,n^{2H-2} \to 0,\end{aligned}$$

as n tends to infinity. Therefore, if $H > 1$,$\rho_H(n) > 0$ for n large enough and $\sum_{n=1}^{\infty}\rho_H(n) = \infty$. We say that the sequence $X_n, n \geq 1$ has *long-range dependence*. Moreover, this sequence presents an aggregation behavior which can be used to describe cluster phenomena. For $H < \frac{1}{2}$, $\rho_H(n) < 0$ for n large enough and $\sum_{n=1}^{\infty}\left|\rho_H(n)\right| < \infty$. In this case, $X_n, n \geq 1$ can be used to model sequences with intermittency.

A fractal model based on fractional Brownian motion (fBm) is a non-stationary model often used to describe random phenomenon. It's named after Robert Brown, regarding his work on the random movement of particles. In a fractional Brownian model (Lehamel and Hammouche, Lehamel and Hammouche), the expected value of the intensity difference between two points is zero but the expected value of the square of the difference is proportional to the distance between the points at a power H where H is the Hurst parameter which is a constant lying in the range$[0,1]$. Brownian motion observed naturally is seen on

a surface with a topological dimension of two. The dimension of the curve would therefore depend on its roughness which is dependent on the parameter H. The Hurst parameter is related to FD by

$$\dim_{fBm} = E + 1 - H \tag{2.19}$$

where E is the *Euclidean dimension* that contains the fBm (i.e. the position of each point of the process is described with the vector $\mathbf{x} = (x_1, \ldots, x_E)$), $\dim_{fBm}$ is the dimension of the Brownian curve. Fractal Brownian functions are statistically self-affine (B. Mandelbrot, 1983). The FD of a fractal Brownian function is invariant with transformations. Several estimation methods of the Hurst parameter are proposed in the fractal literature and where the relationship between the measure N_r and scale r is governed by a powerlaw:

$$N_r \propto r^H$$

2.5.2 Wavelet Based Estimation of the Fractal Dimension

As defined before, for a 2D self similar process, the fractal dimension is Eq. (2.19), where H is the self-similar parameter, known also as Hurst parameter from its introduction (Hurst, 1951). For a 2D fractional Brownian motion (fBm) the Hurst coefficient can be estimated from the slope of the linear equations given by

$$\log_2 E\left[\left|d^i_{j,\mathbf{k}}\right|^2\right] = -(2H+2)\,j + C_i, \tag{2.20}$$

where

$$d^i_{j,\mathbf{k}} = 2^j \int B_H(\mathbf{u})\,\psi^i(2^j\mathbf{u} - \mathbf{k})\,d\mathbf{u}, \tag{2.21}$$

are the wavelet coefficients (for $i = horizontal$, $vertical$, and $diagonal$ directions, for the j scale, and $\mathbf{k}$ translation) of the fBm B_H with ψ denoting the 2D wavelet function, and

$$C_i = \log_2 \frac{\sigma_H^2}{2} V_{\psi^i}(H) \tag{2.22}$$

depending on the variance of the fBm (σ_H^2) and on the wavelet function (see, (Nicolis et al., 2011)). From the slope $\alpha = -(2H+2)$ of the regression in Eq. 2.20 is easy to calculate the Hurst parameter as $H = \frac{-(\alpha+2)}{2}$, and consequently the fractal dimension as in Eq. 2.19.

2.6 Aspects of Texture

Texture is referred as a repetitive visual structure that contains many similar and organised elements that don't stand out and form a global unity impression (Dobrescu and Popescu, 2011). One texture can show off many aspects like brightness, colour, smoothness (fineness), coarseness (roughness), density, uniformity, regularity, linearity, frequency, phase, randomness, granularity and spatial orientation (lay or directionality). In B. Mandelbrot (1983), B. B. Mandelbrot called texture as an "elusive notation" which it is indeed. Thus, a rigorous definition of a texture represented by an image is not an easy task.

Texture analysis will be studied using various approaches, such as lacunarity, grey level co-occurrence matrices (GLCM) and the features extracted from them, box-counting fractal dimension and succolarity proving to be a good choice for texture analysis and classification (Popescu et al., 2017; De Melo and Conci, 2013; Rani and Aggarwal, 2013; N'Diaye et al., 2013; Shahzad et al., 2010).

2.6.1 Lacunarity

Lacunarity is the deviation of a fractal from translational inavariance (Gefen et al., 1983). Translational invariance can, of course, also be a property of nonfractal sets. In addition, translational invariance is highly scale dependent; sets which are heterogeneous at small scales can be quite homogeneous when examined at larger scales or vice versa. Lacunarity can thus be considered a scale-dependent measure of heterogeneity or texture of an object, whether or not it is fractal (Allain and Cloitre, 1991).

The notion of lacunarity makes it possible to distinguish sets that have the same fractal dimension but different textures.

Mandelbrot introduced the lacunarity to further characterize fractal objects (B. Mandelbrot, 1983). It is a measure that describes the distribution of the sizes of gaps or lacunae surrounding the object within the image. Greater lacunarity reflects a greater size distribution of the lacunae or, said in another way, a higher degree of "gappiness". It can be used to distinguish objects with similar fractal dimensions, as some objects with the same or similar fractal dimension will fill the space differently (B. Mandelbrot, 1983), but lacunarity can also be used independently as a general tool for describing spatial patterns (Plotnick et al., 1996).

Several methods have already been introduced in the literature for measuring the lacunarity (B. Mandelbrot, 1983; Gefen et al., 1983; Tolle et al., 2008; Plotnick et al.,

1996), here we have adopted the intuitively clear and computationally simple gliding box algorithm (similar to the box counting algorithm used to estimate fractal dimension) proposed by Allain and Cloitre (1991) and described as follows (Tolle et al., 2008; Gould et al., 2011): One chooses an unit $r \times r$ box of size r, it is placed at the top left corner of an image and counts the number of signal pixels or set points, p, in the box. The box is then glided one pixel to the right and the number of pixels in the $r \times r$ box is counted. This procedure is then repeated until the box reaches the bottom right corner of the image, as the box is centred, in turn, about each point within the set, creating a distribution of box masses $n(p,r)$, the number of boxes with p points and radius r. If l is the size of the image, the number of boxes of width r is $N(r) = (l-r+1)^2$. The width r of the box is chosen such that its numerical value is equal to 2, 4,8, ..., l. For each box width r, the following two quantities are measured: the sum of the number of signal pixels in each box ($Q_1 = \sum_i p\,(i,r)$) and the sum of the square of the number of signal pixels in each box ($Q_2 = \sum_i p\,(i,r)^2$). This distribution is converted into a probability distribution, $Q(p,r)$, by dividing by the total number of boxes $N(r)$ of size r,

$$Q(p,r) = \frac{n(p,r)}{N(r)}$$

The first and second moments of this distribution are now determined:

$$Z_1(r) = \sum_p p\,Q(p,r), \quad Z_2(r) = \sum_p p^2\,Q(p,r)$$

Lacunarity $L(r)$ at a box size r is defined as

$$L(r) = \frac{N(r) \times Q_2}{Q_1^2} = \frac{Z_2(r)}{[Z_1(r)]^2}. \tag{2.23}$$

This calculation is repeated over a range of boxes sizes, ranging from $r = 1$. $p(i,r)$ is the number of data points (signal pixels) in the ith box of radius r and $i \in [1, N(r)]$. In the gliding box algorithm, the box of size r is placed over each pixel in the image and $p(i,r)$ is calculated for each box centered on the ith pixel. Thus, lacunarity measurement adds significantly to the description of an object with a known fractal dimension, in that it describes the empty space around the object, and thus it relates how the object fills space.

2.6.2 Gray Level Co-Occurrence Matrix

In statistical texture analysis, texture features are computed from the statistical distribution of observed combinations of intensities at specified positions relative to each other in the image. According to the number of intensity points (pixels) in each combination, statistics are classified into first-order, second-order and higher-order statistics.

Texture is an important characteristic in the image processing. It implies the idea of uniformity and repeatability in the image to form the surface of the object, and can be considered a measure of the variation of the intensity. According to Haralick (1979), an aspect of texture is concerned with the spatial distribution and spatial dependence among the gray levels in a local area. Herein, fourteen measures of textural features which are derived from the Gray Level Co-ocurrence Matrix (GLCM) have been proposed. This method also known as Grey Tone Spatial Dependency Matrix, is a way of extracting the second order statistical texture features by considering the spatial relationship of pixels and its neighbors.

Nowadays, several applications and approaches in breast cancer are based on texture classification. One of the exceptional characteristic of the image processing to describe texture is that it contains variations of intensities, which form certain repeated patterns. GLCM is one of the best known texture analysis method. It is a matrix where the number of rows and columns is equal to the number of the grey levels, G, in the image. The matrix $P(i, j \mid \Delta x, \Delta y)$ is the relative frequency with which two pixels, separated by a pixel distance $(\Delta x, \Delta y)$ occur within a given neighborhood, one with intensity i and the other with intensity j. Each entry (i, j) in this matrix corresponds to the number of occurrences or intensities occur within a given neighborhood, of the pair of grey levels i and j separated by a pixel distance $(\Delta x, \Delta y)$. The matrix element $P(i, j \mid d, \Theta)$ contains the second order statistical probability values for changes between gray levels i and j at a particular displacement distance d and at a particular angle (Θ). Using a large number of intensity levels G implies storing a lot of temporary data, i.e. a $G \times G$ matrix for each combination of $(\Delta x, \Delta y)$ or (d, Θ).

The properties of GLCM are:

1. GLCM is of square in shape because the reference and neighbouring pixels have same range of values.

2. Number of rows and columns equal to the quantization level of the image.

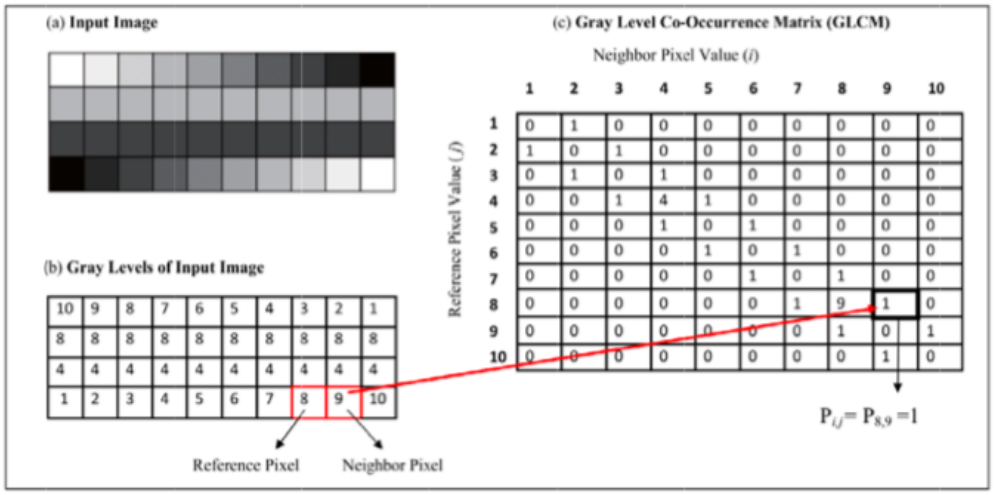


Fig. 2.2 Construction of Gray Level Co-occurrence Matrix (GLCM) from an image.

Khemis et al. (2016) in their work give an overview for calculating fractal attributes based on any statistical features from GLCM on breast tissue characterization. Herein, fractal analysis on the dimensionality of the textural datasets verified that reduced subsets of optimal feature combinations can describe the original feature space adequately for classification. A brief description of this measurements and its formulas are given respectively as follows:

Correlation: Measures the joint probability occurrence of the specified pixel pairs.

$$Correlation = \sum_{i,j} p(i,j) \frac{(i-\mu)(j-\mu)}{\sigma^2} \tag{2.24}$$

Contrast: Measures the local variations present of an image in the GLCM, providing a measure of gray-level contrast between neighboring pixels over the entire image.

$$Contrast = \sum_{i,j} |i-j|^2 \; p(i,j) \tag{2.25}$$

Homogeneity:Measures the closeness of the distributions of elements in the GLCM to the GLCM diagonal. This measure is inversely proportional to GLCM contrast.

$$Homogeneity = \sum_{i,j} \frac{p(i,j)}{1+|i-j|} \tag{2.26}$$

Energy: Also called Angular Second Moment or Uniformity. Provides the sum of squared elements in the GLCM.

$$Energy = \sum_{i,j} p(i,j)^2 . \tag{2.27}$$

2.6.3 Succolarity

Succolarity was defined by Mandelbrot (B. Mandelbrot, 1983) as a parameter which informs about connectivity and inter communication, i.e., indicates the capacity of a flow to cross the set. This definition considers the idea of percolation since percolation means "to flow through", succolare (sub-colare) seems the proper term for "to almost flow through". However, Mandelbrot does not present measures of succolarity in his book. Melo (2007) in his thesis was the first to calculate this fractal measure using a partitioning method adapted from box-counting procedure and extensively studied issues related to estimation of FD, lacunarity and succolarity. This features are likely to be complementary parameters because objects having same porosity and/or FD can differ by the connectedness of their pores or the presence of connecting channels.The notions of succolarity makes possible distinguish different sets or textures that have these same measures.

Differents papers have taken into consideration the succolarity for different real images (De Melo and Conci, 2013; Rani and Aggarwal, 2013; N'Diaye et al., 2013; Shahzad et al., 2010; Popescu et al., 2017).

A Succolarity on fractal sets can be represented as evaluation of the degree of filaments that allow percolation or to flow through according to the direction of entry (Melo, 2007; De Melo and Conci, 2008). This type of analysis is very useful when the input texture has direction and/or flow information associated with it e.g. drainage patterns. The succolarity algorithm revolves on the idea of fractal dimension, based on mathematical and physical properties of the fluid dynamics.

Although De Melo and Conci (2008, 2013) described and explained extensively the method to calculate succolarity using a box counting approach on a square image of side *n*pixels, in Popescu et al. (2017) give the pseudo code for implementing it. This is one of the efficient way for programming. There are some differences between mathematical procedures. One is in the formula for calculating the succolarity. Succolarity algorithm, the most demanding step is the first, do flood input binary image on selected direction (four directions 0^o, 90^o, 180^o, 270^o). This is a classical path finding algorithm. After this, the

next procedures are straightforward. If a pixel that represents empty space (background pixel) on the textured image can be flooded by the fluid then all its four neighbours are examined. This logic is applied recursively on all the pixels of the image until object pixels are encountered. The output of the flooding algorithm is a binary image which exhibits the flooded filaments in the textured image. Then each binary image is partitioned in boxes of size $k \times k$, where $k = 1$ to n and n is the number of the divisions of the image. The occupation percentage of a box, $OP(k)$, is calculated as the ratio between the number of nonobject pixels (background pixels) and the total number of pixels within the box. A pressure field equal to the size of the flooded image is created and consists of linearly increasing of pressure from top to bottom. The pressure at any pc for box size of $k(PR(k,pc))$ is the centroid of the box on the pressure field. Finally, the correlation between occupation percentage and pressure over box is divided by the total flooding case with maximum pressure over box (Popescu et al., 2017):

$$\sigma_{dir}(n) = \frac{\sum_{k=1}^{n} OP(k) \times PR(k,pc)}{\sum_{k=1}^{n} PR(k,pc)} \tag{2.28}$$

where σ represents the succolarity with values in interval $[0,1]$ and *dir* is the chosen flowing direction.

2.7 Machine learning methods

The importance of classifying or discriminating breast lesions between malignant and benign has led many research approaches, among those studies stands out the application of machine learning methods. Maching Learning (ML) have been widely applied in cancer research for the development of predictive models, resulting in effective and accurate decision making. Therefore, these techniques have been utilized with the aim to model the progression and treatment of cancerous conditions. The ability of ML tools to detect key features from complex data sets is one of its most important characteristics.

Supervised learning methods are normally used for classification tasks by using a first data set for training the algorithm and a second data set for validating it. After this two steps the algorithm is ready for implementing the predicting.

In the classification, the models are often evaluated using the Receiver Operating Characteristic (ROC), the Percentage of Correct Classifications (PCC) or accuracy, the sensitivity, and the specificity. ROC analysis provides tools to select possibly optimal models and to discard sub-optimal ones independently from (and prior to specifying) the cost context or the class distribution. A high percentage of PCC, preferably close to 100%, suggests that it is a good classifier (Cortez, 2010). This metric measure can be computed using the equation:

$$\begin{aligned} \Phi(i) = & \begin{cases} 1 & \text{if } y_i = \widehat{y_i} \\ 0 & \text{if } else \end{cases} \\ PCC = & \frac{\Sigma_{i=1}^{N} \Phi(i)}{N} \times 100\% \end{aligned} \tag{2.29}$$

where $\widehat{y_i}$ denotes the predicted value for the $i-$th example.

The basic measures to quantify the diagnostic accuracy of a test include sensitivity and specificity, which are described in terms of TP (Number of true positive assessment), TN (Number of true negative assessment), FN (Number of false negative assessment) and FP (Number of false positive assessment).

- Sensitivity = TP/(TP + FN)
- Specificity = TN/(TN + FP)
- Accuracy = (TN + TP)/(TN+TP+FN+FP)

Sensitivity and specificity are the rates of correctly-assigned positive identifications and negative identifications, respectively.

In this work, to test the efficiency of fractal, lacunarity, succolarity and GLMC features, five techniques (classifiers) of supervised learning for the classification task will be used.

- **Decision Tree:** The Decision Tree (DT) is a branching structure that represents a set of rules, distinctive values in hierarchical form. This representation is analogous to the logical representation IF $-$THEN, which is easy to understand and systematize.
- **Random Forest:** For its part, Random Forest (RF) (Breiman, 2001; Ho, 1995) is an assembly of T Decision Trees without pruning, the nodes of which are split on a

random drawing of m features from the entire feature set M. Each tree is based in a random feature selection from bootstrap sampling tests and RF predictions are built by averaging the outputs of the T trees. This is more difficult to interpret compared to Decision Trees.

- **Neural Networks:** A neural network consists of units (neurons), arranged in layers, which convert an input vector into some output. Each unit takes an input, applies a (often nonlinear) function to it and then passes the output on to the next layer. NN is used for a classification task, especially when the classification component in the network, based on the distribution of the samples, using methods such as k-nearest neighbor. In this work, the NN model is based in the popular multilayer perceptron algorithm, with one hidden layer with H hidden nodes, while the SVM will use a gaussian kernel with one hyperparameter (γ). The advantages of multilayer perceptron are capability to learn non-linear models and capability to learn models in real-time.

- **Support Vector Machines:** Just like the NN model, Support Vector Machines (SVM) is (non)linear function that constructs a separating hyperplane between two classes by maximizing the distance between them, which have also been proposed for data mining tasks (Friedman et al., 2001), obtaining better results when a high nonlinearity is present. Also, NN and SVM are more affected by irrelevant inputs than the DT/RF algorithms, since the latter explicitly perform an internal feature selection (Cortez, 2010).

- **Naive Bayes (NB):** It is used to predict or classify. Although several classification algorithms can use any space of functions, many learning algorithms are probabilistic models taking the form of a conditional or joint probability model. The NB is a predictive classifier, which can be used to take different decision based on the result. NB classification is joint probability model and is based on Bayes theorem with naive (strong) class conditional independence. Class conditional independence means the effect of an attribute value on a given class is independent of the values of other attributes (Friedman et al., 2001).

In this work we will analyze the dependency of the output respect all the set of observed variables, which will be modeled using two supervised learning approaches:

1. Binary Classifier: The abnormality of the lesion in the ROIs is characterized by being: Normal or Abnormal (Benign or Malignant).

2. Muticlass Classifier: There will be 3 levels of classification of the ROIs, Benign (B), Malignant (M) and Normal (N).

We will use a 10-fold cross-validation method to assess the performance of each model (Kohavi et al., 1995; Cortez, 2007). This method partitions the data into 10 subsets of equal size while maintaining the proportionality of group representations. Each model is then trained on 9 subsets (with 90% of the data) and uses the remaining subset to assess its accuracy. This process is then repeated until all subsets have been used as training and test sets. We will then provide the sensitivity, specificity, the area under a ROC curve and the PCC or accuracy for each classifier.

2.8 A Real Data Example

As described before, texture and fractal analysis will be studied using various approaches, such as fractal dimension (box-counting and wavelet spectra), lacunarity, grey level co-occurrence matrices (GLCM) and the features extracted from them, and succolarity proving to be a good choice for texture analysis and classification of the images selected.

In this work, we use a publicly available data from http://web.inf.ufpr.br/vri/breast-cancer-database, called BreaKHis. The BreaKHis database contains microscopic biopsy images of benign and malignant breast tumors. Both breast tumors, benign and malignant, can be sorted into different types based on the aspect of the tumoral cells under the microscope. The dataset currently contains four histological distinct types of benign breast tumors: adenosis (A), fibroadenoma (F), phyllodes tumor (PT), and tubular adenona (TA); and four malignant tumors (breast cancer): ductual carcinoma (DC), lobular carcinoma (LC), mucinous carcinoma (MC) and papillary carcinoma (PC). The distribution of benign and malignant tumors in these classes is presented in Tables 2.1 and 2.2, respectively.

Images were collected through a clinical study from January 2014 to December 2014. It is composed of 7.909 clinically representative from 82 patients using different magnifying factors (40×, 100×, 200×, and 400×) as shown in Figure 2.3. We will only use images with a 400× magnification factor. To date, it contains 588 benign and 1.232 malignant samples. This database has been built in collaboration with the P&D Laboratory [1] –

[1] http://www.prevencaoediagnose.com.br/

Pathological Anatomy and Cytopathology, Parana, Brazil. The main goal is to preserve the original tissue structure and molecular composition, allowing to observe it in a light microscope.

The acquisition of images at different magnifications is performed as follows: first the pathologist identifies the tumor and defines a region of interest (ROI). To cover the whole ROI, several images are captured using the lowest magnification, i.e. 40×. The pathologist preferentially selects images with a single type of tumor (majority of the cases) but some of the images also include transitional tissue, e.g. normal-pathological. In average, a total of 24 images per patient is captured from each slide using the lowest magnification. Then, the magnification is manually increased to 100× and a similar number of images is captured inside the initial ROI. This process is repeated for 200× and 400× magnifications, respectively. A final visual (i.e. manual) inspection discards out-of-focus images.

Table 2.1 Benign image distribution by 400× magnification factor and histological subtypes.

Magnification	**A**	**F**	**TA**	**PT**	**Total**
400×	106	237	115	130	588

Table 2.2 Malignant image distribution by 400× magnification factor and histological subtypes.

Magnification	**DC**	**LC**	**MC**	**PC**	**Total**
400×	788	137	169	138	1.232

2.8.1 Feature extractors and classifiers

Histological tissue images can be characterized by two types of approaches. The first one is based on explicit segmentation to extract structure properties, such as nuclei shape, glandular unit shape, etc., while the second one is a global approach based on texture representation(Spanhol et al., 2016). Since segmentation of histological tissue images is not a trivial task and can be prone to errors, we have chosen a global approach based on fractal and texture representation.

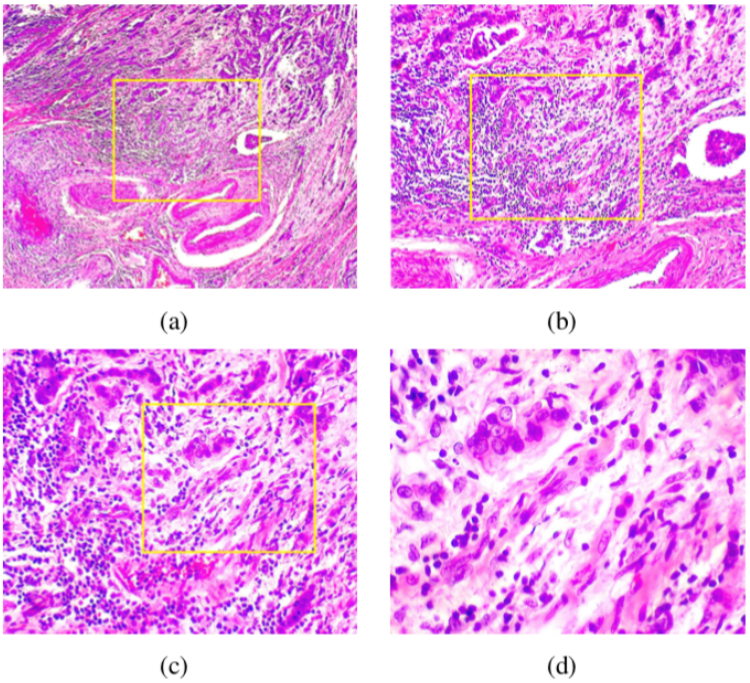
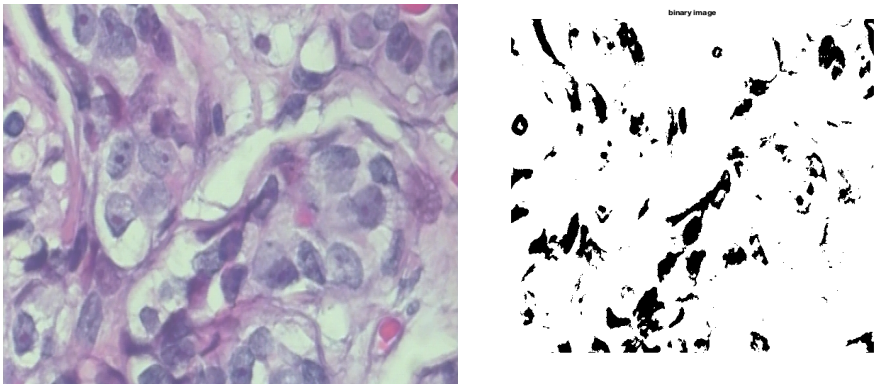


Fig. 2.3 (a) A slide of breast malignant tumor (stained with HE) seen in different magnification factors: (a) 40×, (b) 100×, (c) 200×, and (d) 400×. Highlighted rectangle (manually added for illustrative purposes only) is the area of interest selected by pathologist to be detailed in the next higher magnification factor.

All the statistical analyses reported in this thesis were done using the RMiner (Cortez, 2007, 2016) and RandomForest (RF) (Breiman, 2001; Breiman et al., 2011; Breiman, 2017), both free code libraries in the integrated development environment of the Rstudio (programming language) (Team et al., 2015; R Core Team, 2017). In particular, the library uses the rpart(DT), naivebayes (NB), nnet (NN) and kernlab (SVM) packages.

The images were binarized and resized, in order to facilitate interpretation and features extraction using a series of programs that are freely available from MATLAB. The result was a grey-level image with a resolution of 512×512 pixels by segmentation(Figure 2.4).

A statistical summary of the fractal measures and texture attributes are shown in 2.3. In particular, Table 2.3 shows that the fractal dimension (FD) obtained by the traditional box-counting method presents an average of 1.87 and a minimum of 1.58 over the analyzed images. The maximum of this variable equal to 1.99 corresponds to the mastopathic tissue. Also, a skewness to the left is shown by this measure. An high linear correlation and homogeneity of grey levels are shown by the correlation (Cor) and homogeneity (Hom)



(a) Mastopathic Tissue (Adenosis-A) (b) Binarizated Mastopathic Image

Fig. 2.4 Microscopic biopsy images of benign breast tumors. (a) Mastopathic tissue, in this case adenosis of the breast. (b) Binary images from (a) by segmentation.

Table 2.3 Summary Statistics (Mean, standard deviation, median, minimum, maximum, kurtosis, and skewness) for the following fractal measures and texture features: fractal dimension (FD), entropy (Ent), lacunarity (Lac), wavelet based fractal dimension (FDwav), contrast (Con), correlation (Cor), Homogeneity (Hom), and Energy (Ene).

	Mean.	**Sd.**	**Median**	**Min.**	**Max.**	**Kur.**	**Skew.**
FD	1.87	0.06	1.88	1.58	1.99	3.78	-0.71
Lacunarity	1.49	0.34	1.41	1.03	4.62	16.20	2.56
Contrast	0.05	0.02	0.05	0.00	0.12	3.34	0.28
Homogeneity	0.97	0.01	0.97	0.94	1.00	3.34	-0.28
Correlation	0.89	0.03	0.89	0.76	0.98	3.28	-0.30
Entropy	0.94	0.08	0.97	0.36	1.00	9.56	-2.30
Energy	0.49	0.06	0.47	0.40	0.86	7.64	1.91
Entropy	0.94	0.08	0.97	0.36	1.00	9.56	-2.29
Succolarity	0.66	0.22	0.68	0.10	1.00	1.85	-0.23

indexes, respectively. As expected, the entropy (Ent), energy (Ene) and contrast (con) present lower values.

Figure 2.5 provides the distribution of each attribute by class (cancer and mastopathic images) using a kernel density aproximation. The plot shows that the some attributes such as the FD and Ent present a more accentuated left skewness than other variables. These results are confirmed by the skewness indexes of Table 2.3.

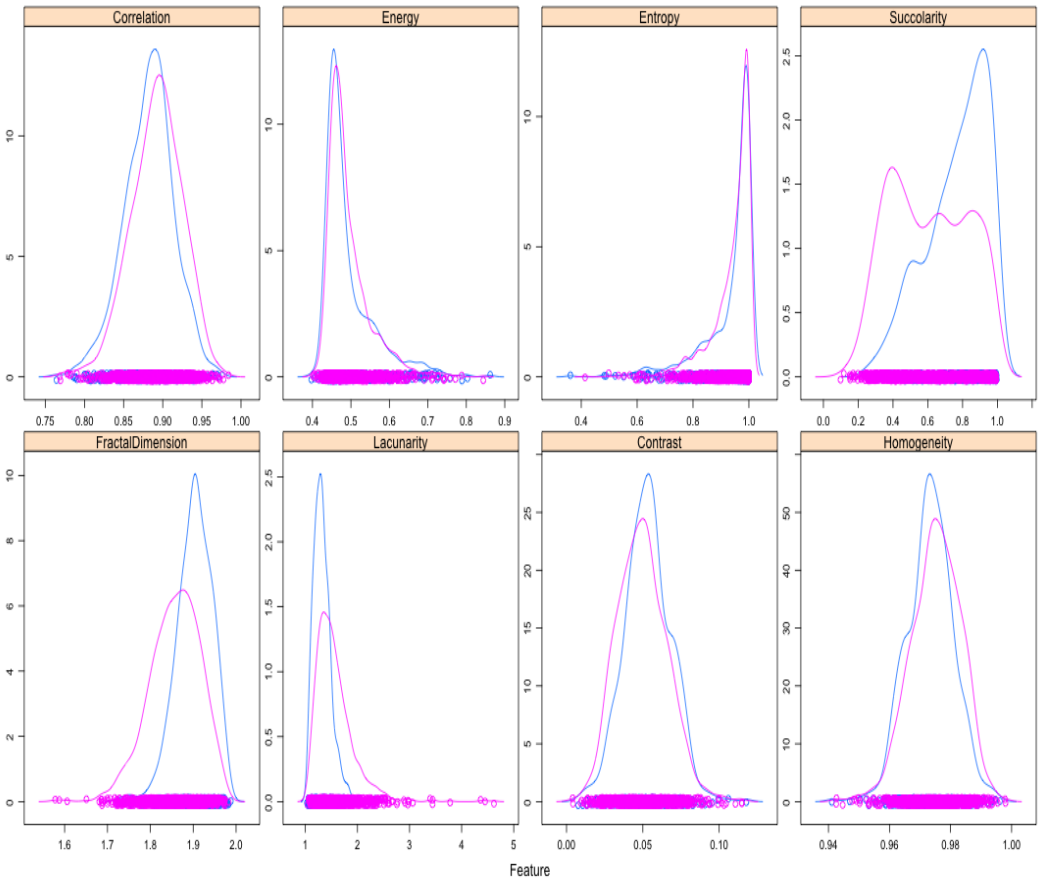


Fig. 2.5 Density plots for each attribute by class in the histological tissue images: the blue lines represent mastopathic images and purple lines represent cancer images.

We then apply some methods (DT, SVM, RF, NB, and NN) for the binary classification (mastopathic versus cancerous) using the co-occurrence matrix (contrast, energy, homogeneity, entropy and correlation), the fractal features (the box-counting and wavelet fractal dimension) and the lacunarity index. The results for the binary and multiclass classification are shown respectively in Table 2.4.

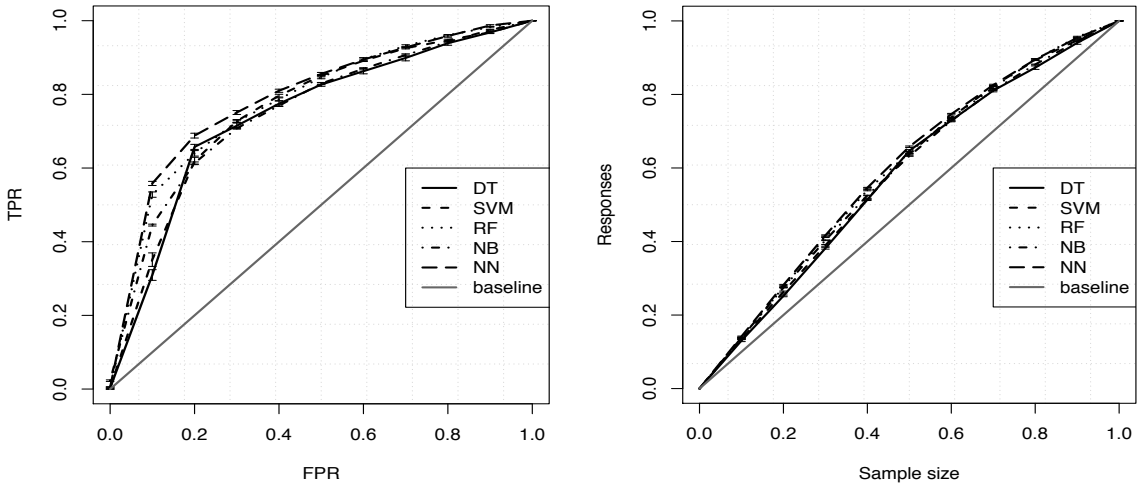
Table 2.4 Means and their 95% *t-Student* confidence intervals of the Area Under a ROC curve (AUC), Accuracy (ACC), and Sensitivity (Sen) for the the binary classification (Normal versus Cancerous) using a 10-fold cross validation.

Models	**AUC**	**ACC.**	**Sen.**	**Sp.**
DT	$0.75_{\pm 0.01}$	$72.42_{\pm 0.31}$	$53.47_{\pm 1.68}$	$81.46_{\pm 0.46}$
SVM	$0.76_{\pm 0.00}$	$73.46_{\pm 0.20}$	$45.88_{\pm 0.34}$	$86.62_{\pm 0.23}$
RF	$0.79_{\pm 0.01}$	$73.46_{\pm 0.28}$	$51.05_{\pm 0.73}$	$84.15_{\pm 0.36}$
NB	$0.76_{\pm 0.01}$	$70.63_{\pm 0.14}$	$68.96_{\pm 0.30}$	$71.43_{\pm 0.16}$
NN	$0.80_{\pm 0.02}$	$74.07_{\pm 0.34}$	$58.38_{\pm 0.67}$	$81.56_{\pm 0.38}$

Based on the results in Table 2.4, the RF and NN gave us the best predictive power. To support the quality of the predictions in the binary classification, in Figure 2.6b the area of the cumulative percentage of captured responses (LIFT accumulative curve) Cortez (2007) is shown for all algorithms used. This is calculated by determining the ratio between the result predicted by our model and the result using no model, for example, taking 60% of the images, and using no model we should get 60% of normal images and using the given model, let's take the SVM algorithm as reference in this case, we should get more than of the 72% of breast cancer images. Sensitivity to the specificity of the cancer class is also plotted using the ROC curve (Fig. 2.6a). These ROC curve suggests a predicative models that are much better than the random classifier baseline (AUC=0.5).

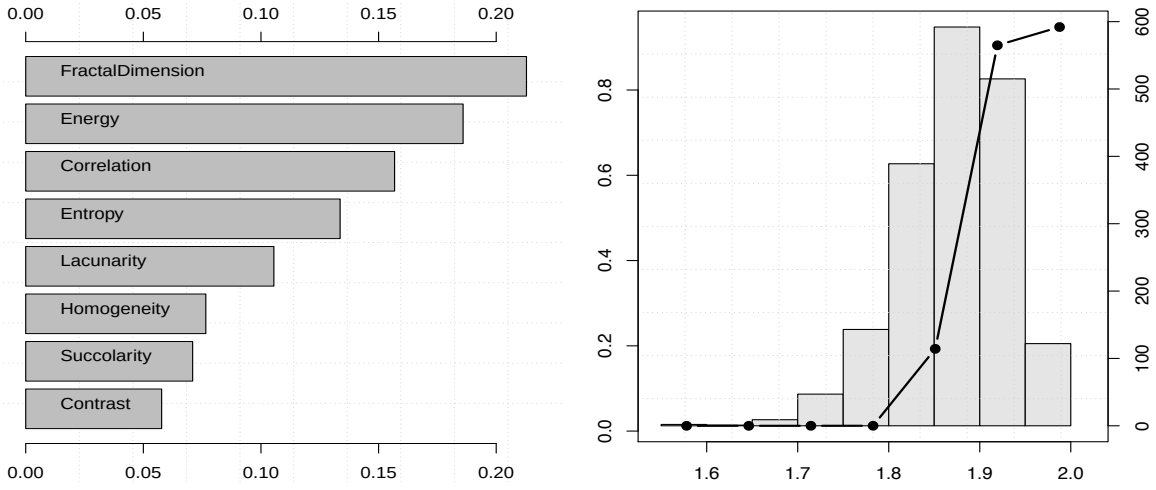
Figure 2.7 present a sensitivity analysis method for extracting input relevance and variable effect characteristic (VEC) curves. The method is described in Cortez (2007), where several sensitivity analysis and visualization techniques are also discussed. Figure 2.7a show the relative importance rate of the inputs and Figure 2.7b the influence of the fractal dimension variable on class mastopathic as measured by the NN algorithm in the binary classification. Here, the same data that was used to fit the model, currently only used to add data histogram to VEC curve.

Particularly, the left of Figure 2.7, confirms the importance of some fractal and texture attributes such as the fractal dimension, energy, correlation, entropy and lacunarity. These five variables present an overall impact of 80% in the NN model, which indicates the importance of combining fractal and lacunarity features with classical measures of the GLMC in the classification of microscopic images of breast tumor tissue images. In the right of Figure 2.7, the VEC curve for Fractal dimension corresponds to the solid line,



(a) (b)

Fig. 2.6 ROC (a) and accumulated LIFT (b) curve for class "Cancer" in binary classification .



(a) (b)

Fig. 2.7 Relative importance of the input variables (a) and variable effect characteristic (VEC) curve for fractal dimension influence on class mastopathic (b) in binary classifications for the NN model.

while the grey bars denote the mastopathic histogram. The points on the graphic means the number of sensitivity analysis levels (e.g. 7). This plot allows to verify that high fractal dimension values are more frequent. Also, a high fractal dimension increases the probability for the mastopathic tissue by more than 0.85 points.

In this work, fractal and texture features obtained from box counting method, wavelet transform, gliding method and GLMC have been analyzed to characterize breast tissues. The study has shown interesting results, the main conclusion is that fractal and texture features combined with machine learning techniques are necessary for explaining the overall complexity of the structure of the breast tumor tissue.

Chapter 3

Multifractal-Based Image Processing

3.1 Introduction

It was shown in Chapter 2 that a fractal object can be characterized by its fractal dimension, which gives a measure of how the density varies with respect to length scale. Fractals reveal how irregular objects can be interpreted and described according to morphological characteristics. The irregular behaviors of these complex structures are difficult or impossible to be quantified by standard modeling techniques or by traditional Euclidean geometry. As we saw in the previous chapters, many of the well-known fractals such as the Sierpinski Carpet are homogeneous since they consist of a geometrical figure repeated on an ever-reduced scale. For these objects, the fractal dimension is the same on all scales. In the real world, fractals are not homogeneous: there is a non-uniformity possessing rich scaling and self-similarity properties that can change from point to point. Put plainly, the object can have different dimensions at different scales. These more complicated objects are known as multifractals, and it is necessary to define a continuous spectrum of dimensions to classify them. An object is considered multifractal when its different regions have different fractal properties.

The concepts underlying the development of multifractals were originally introduced by B. Mandelbrot (1983) and Chhabra and Jensen (1989) in the discussion of turbulence and nonlinear dynamical systems and expanded by Falconer (1986, 1997), Lopes and Betrouni (2009), etc. to many other contexts. Much of the recent interest in multifractals started out with work by Grassberger and Procaccia (1983). A related dimension function was introduced by Radii and Politi (1985). Recently the concepts have been applied successfully

in many different fields including biological systems Nicolis et al. (2017); Derado et al. (2008); Kestener et al. (2001); Stanley (2000) and geophysical systems Posadas et al. (2003); Ferraris et al. (2003); Lovejoy and Schertzer (2007). Some studies have currently characterized some objects as multifractal structures (objects that have different fractal dimensions) instead of monofractal, as regarded in the previously mentioned studies.

In this chapter we look at the use of Multifractals as a tool in image analysis on two free database of images of breast tumor tissue Suckling et al. (1994); Spanhol et al. (2016). We aim to explore the multifractal behavior of different complex breast tissues. Multifractal analysis is a useful way to systematically characterize the spatial heterogeneity of both theoretical and experimental fractal patterns. In this chapter, we use two interesting algorithm for multifractal analysis of breast tissues. This algorithms are based in the computation of the generalized fractal dimensions D_q or moment method and a wavelet based approach and singularity spectrum $f(\alpha)$ for both. Here, the main mathematical methodologies as well as the two practical multifractal metodologies mentioned above are detailed and applied a two microscopic images of breast tumor tissue images.

3.2 Multifractal Analysis

Following Feder (2013), multifractal measures are related to the study of a distribution of physical or other quantities on a geometric support. The support may be an ordinary plane, the surface of a sphere or a volume, or it could itself be a fractal. Then, if a measure has different fractal dimensions on different parts of the support, the measure is a multifractal. Multifractal measures have been observed in many physical situations, for example in fluid turbulence, rainfall distribution, mass distribution across the universe, viscous fingering, neural networks, share prices and in many other phenomena.

There are two basic approaches to multifractal analysis: fine theory, where we examine the structure and dimensions of the fractals that arise themselves, i.e.,we first look at the local limiting behaviour of a measure μ at all points and then globally examine the structure and dimension of the F_α sets partitioned from the local dimension. In the coarse theory on the other hand we only consider the irregularities of the distribution of the measure of balls (or boxes) of small but positive radius (size) r , then take the limit as $r \to 0$. We examine irregularities of the measure at scale r globally and do not look at local behaviour. Since the coarse theory in some way is built on the box-counting dimension and the fine theory on the more precise Hausdorff dimension, the fine theory is more appropriate for

mathematical analysis. When the multifractal spectra should be calculated in practice on the other hand, the coarse theory could be more convenient to use (Falconer, 1997, 2004).

We are going to look at both approaches and ways to calculate multifractal spectra for each theory: moment method and wavelet based approach. From now on we will look at Borel (probability) measures where the domain for the functions are the Borel algebra. The support of the measure will be a bounded region of $\mathbb{R}^n$.

According to Chhabra and Jensen (1989), the $f_H(\alpha)$ singularity spectrum provides a mathematically precise and naturally intuitive description of the multifractal measure in terms of interwoven sets, with singularity strength α, whose Hausdorff dimension is $f_H(\alpha)$.

3.2.1 The fine theory

First of all we need a definition of the local dimension. Take into account definitions 7 and 8 given in chapter 1.

Definition 19. *Given a Borel measure μ and a point x in $spt(\mu)$ the lower and upper local dimensions of μ at x is*

$$\underline{\dim}_{loc}\mu(x) = \liminf_{r\to 0} \frac{\log \mu(B(x,r))}{\log r}$$

and

$$\overline{\dim}_{loc}\mu(x) = \limsup_{r\to 0} \frac{\log \mu(B(x,r))}{\log r}$$

where $B(x,r)$ is a closed ball of radius r centered around x. If these two values are equal then the *common* value is the local dimension of μ at x

$$\dim_{loc}\mu(x) = \lim_{r\to 0} \frac{\log \mu(B(x,r))}{\log r}$$

In some literature the notion of Hölder exponents is used in this case. The exponent in the relationship between $\mu(B(x,r))$ and r for a certain size of r is denoted a coarse Hölder exponent and the limit exponent obtained when r tends to zero is called a local Hölder exponent.

Let μ be a finite Borel measure on $\mathbb{R}^n$. For $\alpha \neq 0$ we now decompose the support of the measure into subsets F_α, where each subset is the set of points where the local dimension is α.

Definition 20. *For $\alpha \neq 0$ define*

$$F_\alpha = \{x \in spt(\mu) : \dim_{loc} \mu(x) = \alpha\}$$

so F_α is the set of points at which the local dimension both exists and equals α.

If the local dimension would not exist we must work with the lower and upper local dimensions.

Definition 21. *For $\alpha \neq 0$ define*

$$\underline{F}_\alpha = \{x \in spt(\mu) : \dim_{loc} \mu(x) \geq \alpha\}$$

and

$$\overline{F}_\alpha = \{x \in spt(\mu) : \dim_{loc} \mu(x) \leq \alpha\}$$

and we will get F_α as the intersection of $\underline{F}_\alpha$ and $\overline{F}_\alpha$.

The next step in the fine theory is to find the dimension of each F_α for $\alpha \neq 0$. In most examples of interest F_α is dense in $spt(\mu)$ for the values of α where F_α is non-trivial, i.e. $F_\alpha = spt(\mu)$. According to equation 2.6, saying that the box dimension can not separate a set from its closure, we get for the lower box dimension

$$\underline{\dim}_B F_\alpha = \underline{\dim}_B \overline{F}_\alpha = \underline{\dim}_B spt(\mu)$$

and similarly for the upper box dimension

$$\overline{\dim}_B F_\alpha = \overline{\dim}_B \overline{F}_\alpha = \overline{\dim}_B spt(\mu)$$

Apparently box dimension is not an appropriate choice of dimension if we want to see differences between the various F_α sets. They will all have the same dimension as the support of the measure. So instead we choose to work with the Hausdorff dimension $\dim H$ and we define (what we will denote the fine spectrum of μ)

Definition 22. *For $\alpha \neq 0$ the Hausdorff (fine) multifractal spectrum of μ is*

$$f_H(\alpha) = \dim_H F_\alpha. \tag{3.1}$$

For the graph of the fine spectrum the following holds.

Theorem 2. *For $\alpha \neq 0$ we have the inequalities*

1. $0 \leq f_H(\alpha) \leq \dim_H spt(\mu)$,
2. $0 \leq f_H(\alpha) \leq \alpha$.

The inequality 2 is derived in e.g. Nordvall (2006), p.76.

Proof. By definition the Hausdorff dimension is non-negative and since it is also monotone any subset of $spt(\mu)$ must have a Hausdorff dimension smaller than that for $spt(\mu)$. □

3.2.2 The coarse theory

In practice it is not possible to estimate the fine spectra based on the Hausdorff dimension from empirical data. Instead a statistical description of how the local dimension varies is often used, captured in the coarse multifractal spectra (Nilsson, 2007).

We now cover the support of the finite measure μ by a $r-$edge and for $r > 0$ and $\alpha \geq 0$ we define

$$N_r(\alpha) = \#\{r-edge\ \ cubes\ \ A\ \ with\ \ \mu(A) \geq r^{\alpha}\}$$

The number of $r-$edge cubes A with

$$r^{\alpha+\varepsilon} \leq \mu(A) < r^{\alpha-\varepsilon}$$

can be expressed by $N_r(\alpha+\varepsilon) - N_r(\alpha-\varepsilon)$.

Falconer (2004) proposes two theories or variations of spectrum of particular interest for mathematicians: The *singularity spectrum or fine theory*, which it was described above, and the *coarse spectrum or coarse theory*, which more adequate for practical purposes. The definitions given below are supported by Falconer (2004), Salat et al. (2017) and Posadas et al. (2003).

Definition 23. *For $\alpha \neq 0$ the coarse multifractal spectrum of μ is*

$$f_C(\alpha) = \lim_{\varepsilon \to 0} \lim_{r \to 0} \frac{\log^+ (N_r(\alpha+\varepsilon) - N_r(\alpha-\varepsilon))}{-\log r} \tag{3.2}$$

if the double limit exists.

Using $\log^+(x) = \max\{0, \log(x)\}$ will ensure that $f_C(\alpha) \geq 0$.

If we assume this limit exists the definition of the coarse spectrum tells us that for $\eta > 0$ and $\varepsilon > 0$ small enough

$$r^{-f_C(\alpha)+\eta} \leq N_r(\alpha+\varepsilon) - N_r(\alpha-\varepsilon) \leq r^{-f_C(\alpha)-\eta}$$

for all sufficiently small r. What this means, roughly speaking, is that the number of $r-$edge cubes with $\mu(A)$ approximately r^{α} obey a power law as $r \to 0$ and that the power law exponent is $-f_C(\alpha)$. That is

$$N_r(\alpha+\varepsilon) - N_r(\alpha-\varepsilon) \sim r^{-f_C(\alpha)}$$

If the limit as $r \to 0$ in the definition of $f_C(\alpha)$ fail to exist we define the following.

Definition 24. *For $\alpha \neq 0$ the lower and upper coarse multifractal spectra of μ are*

$$\underline{f}_C(\alpha) = \lim_{\varepsilon \to 0} \liminf_{r \to 0} \frac{\log^+(N_r(\alpha+\varepsilon) - N_r(\alpha-\varepsilon))}{-\log r}$$

and

$$\overline{f}_C(\alpha) = \lim_{\varepsilon \to 0} \limsup_{r \to 0} \frac{\log^+(N_r(\alpha+\varepsilon) - N_r(\alpha-\varepsilon))}{-\log r}.$$

The limit as $\varepsilon \to 0$ will exist since it is the limit of a decreasing (non-negative) function.

The relationship between the fine and the coarse spectra is given in the following theorem.

Theorem 3. *Let μ be a finite measure on $\mathbb{R}^n$. For $\alpha \neq 0$*

$$f_H(\alpha) \leq \underline{f}_C(\alpha) \leq \overline{f}_C(\alpha)$$

The proof can be found in Falconer (1997), p.188.

Notice the similarity with Equation 2.8 and that just as certain sets can have equal Hausdorff and box-counting dimensions, certain measures have the same fine and coarse spectra (Falconer, 1997).

3.3 Moment Sums and Legendre Transformations

Many measures have spectra that are equal to the Legendre transformation of an auxillary function β based on moment sums. This provides an alternative way of calculating

the coarse multifractal spectra that avoids the often slow and tedious work of directly estimating the power law behaviour of $N_r(\alpha+\varepsilon)-N_r(\alpha-\varepsilon)$.

Given a measure μ, the support is covered by a $r-$edge and for $q \in \mathbb{R}$ and $r>0$ we consider the $q-$th power moment sums

$$M_r(q) = \sum_i \mu(A)^q = \sum_i [P_i(r)]^q \tag{3.3}$$

taken over all cubes A with $\mu(A)>0$.

We identify the power law behaviour of $M_r(q)$ by the functions

$$\underline{\beta}(q) = \liminf_{r\to 0} \frac{\log M_r(q)}{-\log r} \tag{3.4}$$

and

$$\overline{\beta}(q) = \limsup_{r\to 0} \frac{\log M_r(q)}{-\log r} \tag{3.5}$$

When they are equal the auxillary function $\beta : \mathbb{R} \to \mathbb{R}$ is defined by their common value

$$\beta(q) = \lim_{r\to 0} \frac{\log M_r(q)}{-\log r} \tag{3.6}$$

that is the power law exponent in

$$M_r(q) \sim r^{-\beta(q)} \tag{3.7}$$

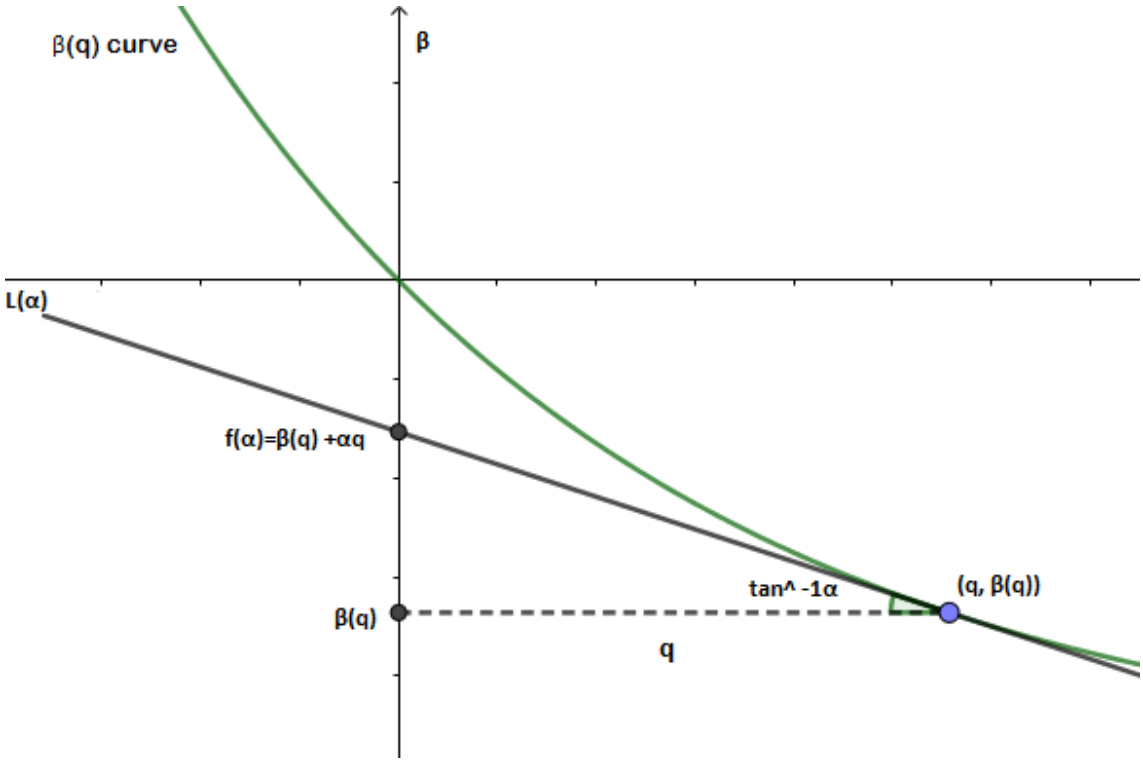


Fig. 3.1 Illustration of the Legendre transform. The figure shows the $\beta(q)$ function plotted against q and the tangent to this curve with slope $-\alpha$. For a given $\alpha(q)$ we receive the Legendre transform of $\beta(q)$ as the intersection of the tangent with the vertical axis.

In order for the Legendre transform of the $\beta(q)$ function to be easily and uniquely determined we need $\beta(q)$ to be a convex, decreasing function. That $\beta(q)$ is convex roughly means that every cord of the graph of $\beta(q)$ lie above or on the graph. Inspecting Figure 3.1 we can see that a line with slope $-\alpha$, through some point $(q^*, \beta(q)^*)$ on the graph of $\beta(q)$, will intersect the vertical axis in $(0, \beta(q^*) + \alpha q^*)$. We receive the Legendre transform of β for a given α in $[\alpha_{\min} - \alpha_{max}]$ as the minimum value of $\beta(q) + \alpha q$ seen as a function of q. Since $\beta(q)$ is convex the minimum value of $\beta(q) + \alpha q$ will be obtained when the line with slope $-\alpha$ is the tangent to the graph, i.e. where $\alpha = -\beta'(q_0)$ for some q_0 in $\mathbb{R}$. More information about the Legendre transform can be found in e.g. Arnol'd (2013) and Courant and Hilbert (1954).

For $\alpha \neq 0$ the upper and lower Legendre spectra of the measure μ can then be defined as the Legendre transforms

$$\underline{f}_L(\alpha) = \inf\left\{\underline{\beta}(q) + \alpha q\right\}$$

and

$$\overline{f}_L(\alpha) = \inf\left\{\overline{\beta}(q) + \alpha q\right\}$$

.

If these values are equal, their common value is the Legendre spectrum of μ.

$$f_L(\alpha) = \inf\{\beta(q) + \alpha q\} \tag{3.8}$$

The following theorem shows how the Legendre spectrum is related to the coarse multifractal spectrum.

Theorem 8. *Let μ be a finite measure on $\mathbb{R}^n$. For $\alpha \neq 0$*

$$\underline{f}_C(\alpha) \leq \underline{f}_L(\alpha)$$

and

$$\overline{f}_C(\alpha) \leq \overline{f}_L(\alpha)$$

This theorem and its proof can be found in Falconer (1997), p.190.

For many measures the Legendre spectrum equals the coarse multifractal spectrum and the variables q and $\beta(q)$ can be used to find α and $f_C(\alpha)$. This approach can have a problem for negative values of q. If a cube only clips the edge of the $spt(\mu)$ then $\mu(A)^q$ can become very large. However there are ways to handle this problem, for example by restricting the sums to cubes with a central portion intersecting the support of the measure (Falconer, 1997).

3.3.1 Generalized correlation dimension function D_q

We consider the support of a multifractal as a $n-$dimensional space which is divided into $n-$dimensional boxes of size r. Then

$$P_i(r) = \sum_{i=1}^{N(r)} \mu(r)$$

denotes the integrated measure on the $i-$th cube of edge r. One can then define the $qth-$order moment of the probability distribution or partition function $M_r(q)$ as described by the equation 3.3, where $N(r)$ is the number of boxes of size r. In the Chapter 2 we saw

that the information dimension (Equation 2.9) and the correlation dimension (Equation 2.12) can then be written from Equation 2.13 as

$$S_q(r) = \frac{1}{q-1} \log M_r(q) \tag{3.9}$$

and D_q, the generalized correlation dimension by Equation 2.14, 3.6 and Equation 3.9 as

$$D_q = \frac{1}{q-1} \lim_{r \to 0} \frac{\log M_q(r)}{\log\left(\frac{1}{r}\right)} = \frac{\beta(q)}{q-1} \tag{3.10}$$

The multifractal structure is characterized by obtaining the generalized dimension D_q, which is related to a value of q. Variable q is the exponent that expresses the fractal properties in different scales to an object, ranging between $-\infty$ and $+\infty$. The plot of D_q versus q is generally sigmoidal and decreasing to a multifractal structure. Some studies consider determined values D_q that is D_0, D_1 and D_2, which depict the multifractality of an object when the inequality condition $D_0 \geq D_1 \geq D_2$ is satisfied. Values D_q depict the multifractal object that can be compared by the single fractal dimension method, so the generalized dimensions for D_0, D_1 and D_2 are known as the box counting dimension (also capacity dimension) D_0, entropy dimension D_1, and correlation dimension D_2, respectively (Rényi, 1955), which are then given as

$$\dim_B = \lim_{q \to 0} D_q = D_0$$

$$\dim_I = \lim_{q \to 1} D_q = D_1$$

$$\dim_C = \lim_{q \to 2} D_q = D_2$$

The generalized dimension function D_q is defined for all real values of q and is estimated as the slope of the $\log - \log$ plot of $S_q(r)$ vs r. For monofractal sets, the function D_q is a linear function of q; in order words, no additional information is gained by examining higher moments, that is, more extreme values of the measure μ. For a simple fractal or a monofractal structure, D_q is independent of the value of q. Alternatively, for multifractal sets, D_q is a nonlinear function of q. Note that there are lower and upper limiting dimensions, $D_{-\infty}$ and $D_{+\infty}$, which are related to the regions of the set where the measure μ is sparser and denser, respectively. For positive values of q, D_q reflects the

scaling of the large fluctuations and strong singularities. In contrast, for negative values of q, D_q reflects the scaling of the small fluctuations and weak singularities (Vicsek, 1993).

3.3.2 Mass exponent function $\beta(q)$

For multifractal measurements and a given value of r, a probability distribution is measured

$$\sum_i [P_i(r)]^q \sim r^{-\beta(q)} \tag{3.11}$$

where $\beta(q) = (q-1)D_q$ is a mass exponent function, $q \in (+\infty, -\infty)$, and $i \in [1, N(r)]$ (Plotnick et al., 1996). The numerical values of D_q corresponding to different values of q are the generalized fractal dimensions of a structure (see Equation 2.14). For monofractal sets, $\beta(q)$ is linear, $\beta(q) = qH - 1$, where H is the *Hurst exponent*. In contrast, multifractal sets display a nonlinear function $\beta(q)$. Moreover, for self-similar multifractals, Equation (3.9), (3.10), (3.7) , (3.11) lead to

$$\beta(q) = (q-1) D_q = \lim_{r \to 0} \frac{\log M_q(r)}{\log(r)} \tag{3.12}$$

or $\beta(q)$ can be defined as:

$$\beta(q) = \lim_{r \to 0} \frac{\left[\log \left(\sum_i P_i(r)^q \right) \right]}{\log\left(\frac{1}{r}\right)} \tag{3.13}$$

where $P_i(r) = \frac{M_i}{M_0}$ is density, M_i is the number of pixels within the ith box and M_0 is the number of pixels for all image. $\sum_i P_i$ is the density for all boxes (i) at a determined scale r.

Note that for $q = 0$, $D_q = \beta(q) = D(0)$; and $\beta(1) = 0$.

3.3.3 Multifractal spectrum $f(\alpha)$

Another way to calculate the multifractal spectra is through the relationship between parameters $f(\alpha)$ versus α, where

$$N(\alpha) \sim r^{-f(\alpha)} \tag{3.14}$$

$N(\alpha)$ is the number of boxes for which probability $P_i(r)$ of finding a pixel within a given region i scales as

$$P_i \sim r^{\alpha_i}$$

$f(\alpha)$ is the fractal dimension to the all regions with singularity strengths between α and $\alpha + \delta\alpha$, where α takes on values between $-\infty$ and $+\infty$. The number $\alpha_i = \log \mu_i(r)/\log r$, also referred to as the *Hölder exponent*, is the singularity strength of the ith box. This exponent may be interpreted as a crowding index of a measure of the concentration of measure μ: the greaterα_i is, the smaller the concentration of the measure, and vice versa. For every box size r, the numbers of cells $N_\alpha(r)$ in which the Hölder exponent α_i has a value within the range $[\alpha, \alpha + \delta\alpha]$ behave like Equation (3.14).

The function $f(\alpha)$ signifies the *Hausdorff dimension* of the subset which has singularity α_i equal to α; that is, $f(\alpha)$ characterizes the abundance of cells with Hölder exponent α and is called the singularity spectrum of the measure. The measure μ is said to be a *multifractal measure* if its singularity spectrum $f(\alpha) \neq 0$ for a range of values of α.

Theiler (1990) defined a description of the generalized fractal dimension in a manner analogous to the box-counting dimension such that let $\beta(q) = (q-1)D_q$, the singularity spectrum $f(\alpha)$ and the mass exponent function $\beta(q)$ are connected via the Legendre transform from the variables (q, β) into a new set of variables (α, f) as

$$\alpha(q) = \frac{\partial \beta}{\partial q} \qquad f(\alpha) = q\alpha(q) - \beta(q) \tag{3.15}$$

and

$$q = \frac{\partial f}{\partial \alpha} \qquad \beta(q) = q\alpha(q) - f(\alpha) \tag{3.16}$$

The $f(\alpha)$ spectrum and the generalized dimension function D_q contain the same information, both characterizing the interwoven ensemble of fractal dimensions $f(\alpha)$. To calculate $f(\alpha)$, two functions $f(q)$ and $\alpha(q)$ should be calculated. The functions $f(q)$ and $\alpha(q)$ are defined as follow:

$$f(q) = \lim_{r \to 0} \frac{\sum_i \mu_i(q,r) \log(\mu_i(q,r))}{\log r} \tag{3.17}$$

$$\alpha(q) = \lim_{r \to 0} \frac{\sum_i \mu_i(q,r) \log(P_i(q,r))}{\log r} \tag{3.18}$$

where $\mu_i(q,r) = \frac{\mu_i(r)^q}{\sum_{i=1}^{N(r)} \mu_i(r)^q} = \frac{[P_i(r)]^q}{\sum_i [P_i(r)]^q}$, $i \in [1, N(r)]$ (Posadas et al., 2003).

$f(q)$ and $\alpha(q)$ are functions of the moments of the distribution. $f(q)$ is the slope of the line obtained by plotting $(\sum_i \mu_i(q,r) \log(\mu_i(q,r)))$ against $\log r$, and $\alpha(q)$ is the slope of the line obtained by plotting $(\sum_i \mu_i(q,r) \log(P_i(q,r)))$ against $logr$. The values of these functions for each order of the moment q allow us to construct the multifractal spectrum. Chhabra and Jensen (1989) were the first to suggest and demonstrate that these formulas provide a practical, efficient, and highly accurate method for the direct computation of singularity spectrum $f(\alpha)$.

The parameter q provides a scanning tool to scrutinize the denser and rarer regions of the measure μ. For $q > 1$, regions whereμ has a high degree of concentration are amplified, while for $q < -1$, regions with a small degree of concentration are magnified. Finally, for $q = 1$, the measure itself is replicated. The function $f(\alpha(q))$ thus gives the entropy dimension of the distorted measure $\mu(q,r)$ and characterizes the original measure μ by analysing the variation under successive distortions driven by the parameter q. The singularity spectrum $f(\alpha(q))$ takes its maximum value for $q = 0$ and typically has a parabolic shape around this point. The number $f(\alpha(0)) = \alpha(0) = D(0)$ is the box dimension (or fractal dimension) of the measure μ, and the number $f(\alpha(1)) = \alpha(1) = D(1)$ is the information dimension.

For a multifractal object, plot $\alpha(q)$ versus $f(\alpha(q))$ fits a parabola with concavity turned down. From values α, it was possible to calculate the extension of singularity length $\Delta\alpha$ and the curve asymmetry of singularity spectrum. Extension of singularity length $\Delta\alpha$ is defined as follows:

$$\Delta\alpha = \alpha_{\max} - \alpha_{\min}$$

$\alpha_{\max}$ (highest value of α) and $\alpha_{\min}$ (lowest value of α) indicate the fluctuation of minimum and maximum probability of pixels, respectively. The higher the $\Delta\alpha$, the larger the probability distribution; furthermore, the multifractality is stronger and the pixel distribution of image is more complex (Hu et al., 2009).

The curve asymmetry of singularity spectrum (A) was calculated according to the following expression:

$$A = \frac{\alpha_0 - \alpha_{\min}}{\alpha_{\max} - \alpha_0}$$

The curve of singularity spectrum is symmetric to $A = 1$, the curve is left-skewed if $A > 1$ and is right-skewed if $A < 1$. When the spectrum is left-skewed, it means that there

is stronger presence of high fractal exponents and significant fluctuation; otherwise, it indicates the domain of low exponents and slight fluctuation (Hu et al., 2009; Costa and Nogueira, 2015).

3.4 Wavelet-based multifractal spectrum (WMFS)

Several methods have been developed to study the multifractality of measures in different contexts. Although all of them are guided by the same principle of assessing fractal dimensions of self-similar structures with varying regularities and to produce the distribution of indices of regularity, which constitutes the multifractal spectrum (MFS). The multifractal formalism relates the MFS to the partition function measuring high-order dependencies in the data. In this section an interesting and robust method for estimating the multifractal wavelet spectrum is analyzed. The main approaches of Stehlık et al. (2016), Nicolis et al. (2017), Nicolis et al. (2011) and Derado et al. (2008) are used. Here, multifractal analysis is concerned with describing the local singular behavior of measures or functions in a geometrical and statistical fashion.

In Stehlık et al. (2016), the fractal hypothesis in cancer research is discussed for case of mammary cancer. They conducted a simple discrimination on the basis of box-counting dimension. Moreover, they discuss on statistical distributions of fractal dimensions for both mammary cancer and mastopathy, showing a good fit for gamma distribution. They detected significant differences in the underlying distribution between the two groups. A multifractal analysis on the basis of a wavelet based approach has been conducted. Discussion on alternative cancer therapy and cancer prevention is provided.

In Nicolis et al. (2017), the discussion of necessity for multifractal measures is provided. To support the diagnostics, we shall be aware of this differentiation and we shall use robust methods for measuring of multifractality. The main reason why robust discrimination is needed is the simple fact that histopathological discrimination between mastopathy and cancer is difficult in many cases. As authors show in this article, using spectra or energy for multifractal model can be of help. However, it needs robustness since many mastopathic tissues are mimicking cancer ones. This article tackles this multidisciplinary topic.

In the following we will describe the wavelet-based multifractal spectrum (WMFS) proposed by Riedi (1999, 1997); Nicolis et al. (2017) and we will apply it to a two microscopic images of breast tumor tissue images (cancer and mastopathic tissue). Wavelet transform is another method originally implemented for time series analysis (Ramirez-

Cobo et al., 2011; Lu and Berliner, 1999), then suited for a generalization to higher dimensions (see, Derado et al. (2007, 2008), (Ramirez-Cobo et al., 2011; Nicolis et al., 2011)). The advantages of using the wavelet-based MFS are availability of fast algorithms for wavelet transform, the locality of wavelet representations in both time and scale, and intrinsic dyadic self-similarity of basis functions.

The multifractal formalism to nonconservative and continuous phenomena is based on the concepts of the partition function and the Legendre transform. The partition function, $T(q)$ can be defined in terms of wavelet coefficients as

$$T(q) = \lim_{j \to -\infty} \log_2 E|d_{j,k}|^q \tag{3.19}$$

where $d_{j,k}$ is the wavelet coefficient at level j and location k, and q is the order of moments. We emphasize that q is a real number within a certain range covering the negative numbers as well (Derado et al., 2008).

Even though Eq. (3.19) is very informative, the singularity measure is not explicit. Derado et al. (2008) proposed that the local singularity strength could be measured in terms wavelet coefficients as :

$$\alpha(t) = \lim_{k2^j \to t} \frac{1}{j} \log_2 |d_{j,k}| \tag{3.20}$$

where $d_{j,k}$ is the normalized wavelet coefficient at scale j and location k. The local singularity strength measure Eq. (3.20) converges to the local Hölder index the process at time t. Small values of $\alpha(t)$ reflect the more irregular behavior at time t. Any inhomogeneous process has a collection of local singularity strength measures and their distribution $f(\alpha)$ forms the MFS. A direct way to obtain this spectrum is to use the counting technique, that is,

$$f(\alpha) = \lim_{\varepsilon \to 0} \#\{\alpha(t) : \alpha - \varepsilon < \alpha(t) < \alpha + \varepsilon, \quad -\infty < t < \infty\}. \tag{3.21}$$

Although it is feasible to estimate the multifractal spectrum using Eq. (3.20) and Eq. (3.21), the method is not practicable due to the difficulty of approximating the limit as well as the large computational complexity. A useful tool to make estimation efficient is the Legendre transform. The Legendre transform of the partition function is defined as

$$f_L(\alpha) = \inf_q \{q\alpha - T(q)\}. \tag{3.22}$$

It can be shown that $f_L(\alpha)$ converges to the true MFS using the theory of large deviations.

If we rearrange Eq. (3.19), it becomes

$$E|d_{j,k}|^q \sim 2^{jT(q)} \quad \text{as} \quad j \to -\infty. \tag{3.23}$$

A standard linear regression can be used to estimate the the partition function $T(q)$ since the values $E|d_{j,k}|^q$ could be easily obtained by moment-matching method. Formally,

$$\log_2 \widehat{S_j(q)} = jT(q) + \varepsilon_j, \tag{3.24}$$

where $\widehat{S_j(q)} = \frac{1}{2^j}\sum_{k=1}^{N2^{-j}} |d_{j,k}|^q$ is the empirical q^{th} moment of the wavelet coefficients (N is the length of the time series) and the error term ε_j is introduced from the moment matching method when replacing the true moments with the empirical ones.

The ordinary least square (OLS) estimator gives the estimation of the partition function,

$$\widehat{T}(q) := \sum_{j=j_1}^{j_2} a_j \log_2 \widehat{S}_j(q), \tag{3.25}$$

where the regression weights a_j must verify the two conditions $\sum_j a_j = 0$ and $\sum_j j a_j = 1$ (Delbeke and Abry (Veitch and Abry, 1999)). Thus, we can estimate $f(\alpha)$ though a local slope of $\widehat{T}(q)$ at values

$$\widehat{\alpha}(q_l) = [\widehat{T}(q_{l+1}) - \widehat{T}(q_l)]/q_0, \quad q_l = lq_0$$

as

$$\widehat{f}(\alpha(q_l)) = q_l\alpha(q_l) - \widehat{T}(q_l).$$

Multifractal spectra can be found even for monofractal processes, the spectra generated from monofractal processes are ramp-like with a dominant (modal) irregularity corresponding to the theoretical Hurst exponent (see Riedi et al. (1999)). Figure 3.2 compare the multifractal spectrum for a monofractal signal and a multifractal one.

3.4.1 Multifractal descriptors

The multifractal spectrum can be approximately described by three canonical descriptors(Stehlık et al., 2016) , which are:

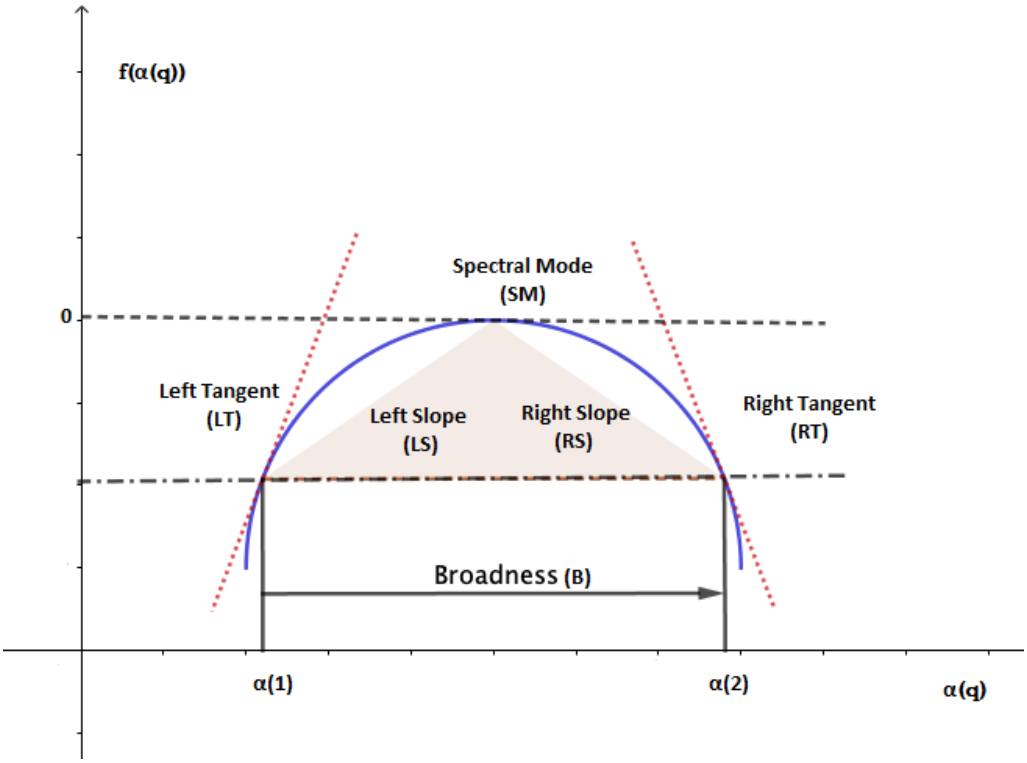


Fig. 3.2 Illustration of geometric descriptors of multifractal spectra. Note that horizontal axis represents values of Hölder regularity index $\alpha(q)$, while vertical axis represents values proportional to the relative frequency of these indices, $f(\alpha(q))$.

1. **Spectral Mode** (Hurst exponent, *SM*): Represents the apex of spectrum or most common Hölder regularity index α found within the signal,

2. **Left slope** (*LS*) or **left tangent** (*LT*): Represents the slope of the distribution produced by the collection of Hölder regularity index α with smaller values of the mode (*SM*), and

3. **Width spread** (Broadness, *B*) or **right slope** (*LS*) or **right tangent** (*RT*): The broadness (*B*) is more intricate descriptor of the multifractal spectrum. Broadness (*B*) is believed to be a more meaningful than right slope (*RS*) or right tangent (*RT*) because it is a compound measure representing the overall nature of the multifractal spectra, taking into account the overall variability among the Hölder regularity index α.

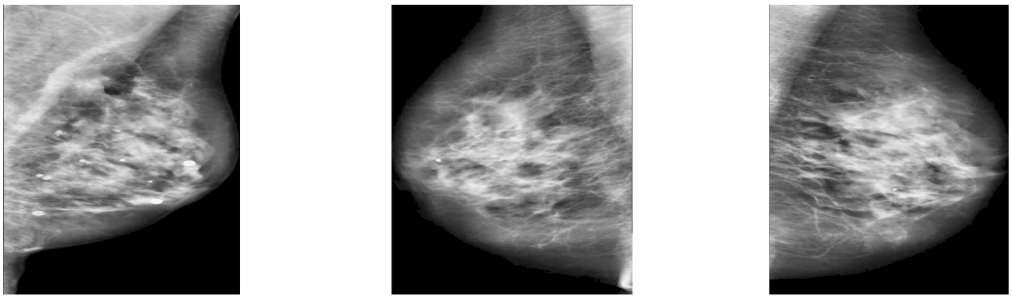
The overall multifractal descriptors are also graphically presented in Figure 3.2. In addition, broadness (B) partially accounts for right slope (RS) or right tangent (RT) in calculation, as the resultant value of B is based on the relative values of RS and LS. The both slopes (or both tangents) can be obtained easily using the interpolation technique, while it is not straightforward to define the broadness (B) automatically. There are many ways to define the broadness (B). In this work, we select the method proposed by Stehlık et al. (2016).

3.5 Application to Two Real Datasets (Suckling et al., 1994; Spanhol et al., 2016)

To illustrate our methods we examine two different situations. First, we provide an application of the multifractal analysis based on generalized fractal dimension or moment method of three examples of benign, malignant and normal mammograms from MIAS database (Suckling et al., 1994) shown in Figure 3.3. Also, we analyze this approach on two histopathological images of breast tumor tissue (benign and malignant) using 400× magnifying factor from Spanhol et al. (2016) in conjunction with the wavelet spectrum-based approach. Herein, the MATLAB program (Vadakkan, 2009; Costa, 2013) was used to calculate the Fractal and Multifractal measures from binary images, specifically the functions $f(q)$ and $\alpha(q)$ using the equations above. It was implemented and used to calculate the multifractal spectra based on the method and ideas proposed by Posadas et al. (2003). The computing time in 2 dimensional models is very low, so it can give a first measure about the multifractal structure of the samples.

3.5.1 Mammographic Image Analysis Society database (MIAS) (Suckling et al., 1994)

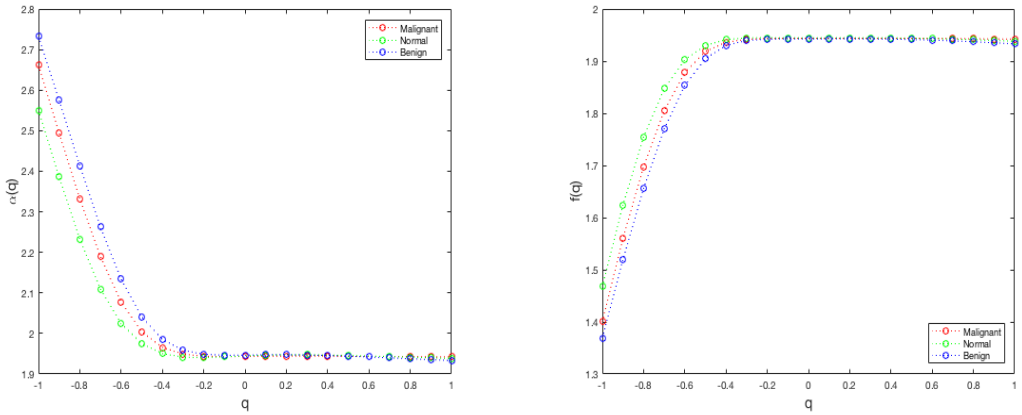
Here we have slices of malignant, normal and benign mammograms of 1024×348 pixeles (Fig. 3.3a), 1024×459 pixeles (Fig. 3.3b) and 1024×451 pixeles(Fig. 3.3c) respectively, so for the box counting method they have been resized to $512 \times 512 pixeles$. For it images examined using the multifractal approaches, the two functions $\alpha(q)$ and $f(q)$ for each q were determined from slope of plots of the numerators of Eq. 3.17 and 3.18 *versus* $\log r$ over the entire range of r values considered (2-250 pixels). The range of q values



(a) (b) (c)

Fig. 3.3 Examples of Malignant, Normal and Benign mammographic images selected for the multifractal analysis.

over which both functions were linear, q was selected considering the coefficients of determination (R^2) of both fits (Fig. 3.4 and 3.6). Note that, in our study the range of q over which both functions were computing take values in the interval $[-1,1]$. The $f(q)$ and $\alpha(q)$ functions obtained over a given q (Fig.3.4), were used to construct the $f(\alpha)-$spectra as an implicit function of q and r. The symmetry of multifractal spectra was evaluated by comparing the width of the spectra from their center $\alpha(0)$ to $|q_i|$. Values of $|q_i|$ were the same in both the positive and negative domains equal to the smaller of the two defining a Δq interval. The largest variation in Δq was observed in the range of negative of q values (Fig.3.4 and Fig.3.6).



(a) (b)

Fig. 3.4 Multifractal spectrum $\alpha(q)$ (a) and *f(q)* (b) for each order of the moment q for particular mammography images (Normal, Benign and Malignant).

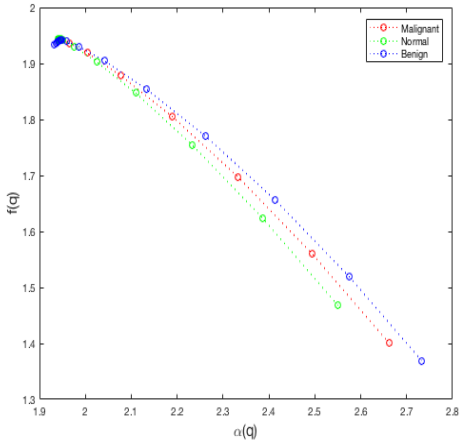
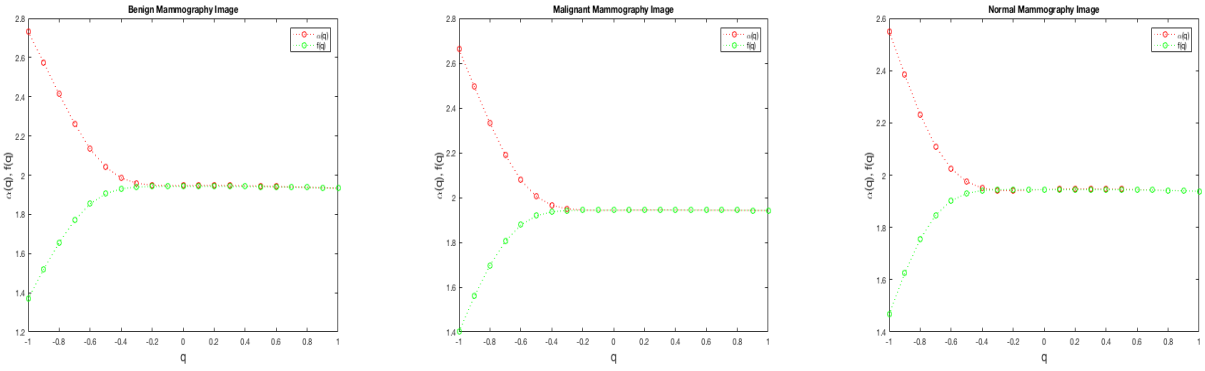


Fig. 3.5 The multifractal spectra of the binary images of the Malignant, Benign and Normal breast cancer.

The multifractal spectrum was plotted to visualize the distribution of the space occupied by breast cancer structures. The multifractal spectra corresponding to only three mammogram images divided into three groups according to severity of abnormality are shown in Fig. 3.5. The crest value of the multifractal spectra corresponds to D_0. The shift in the crest of the spectra from top to bottom is due to differences in the values of D_0, corresponding to the apex of the spectrum. Such as the Benign mammographic image occupy more space than the normal and malignant mammographic images, and hence their D_0 values were significantly similar. The normal mammographic image occupy smallest space with the lowest $\alpha(q)$ values.

The breast cancer structure is only considered a multifractal structure if there is a statistically significant difference between D_0, D_1, and D_2 (see Figs. 3.4 and 3.5).

Similarly, a large value for $\alpha_{max} - \alpha_{min}$ of examples malignant, normal and benign mammograms, and thus the wide opening of the curve, indicated the multifractal nature of these properties (Fig. 3.4 to Fig. 3.6). Despite the nearby values, their multifractal spectrum gives us a complete information in the hierachical breast cancer structure for the selected images. Benign shown a weak multifractal character in concordance with previous results. Normal presented a strong sclaing in the low values, being stronger that malignant. The changes discussed above in the same spectra for selected images revealed significant differences among the different structures of breast cancer.



(a) (b) (c)

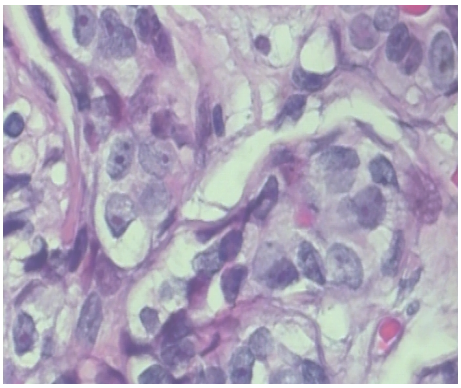
Fig. 3.6 Multifractal analysis of three different textures of mammography images (Benign, Malignant, Normal).

3.5.2 Histopathological Images (Spanhol et al., 2016)

In this section we are going to look at the results received from running the implemented moment method on two types of images of breast tissue., i.e. this procedure was applied on one image of mammary cancer tissue and one image of masthopatic tissue.

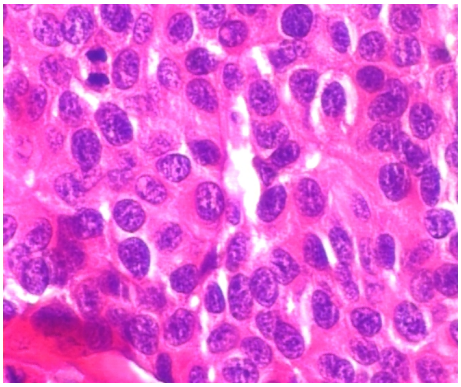
The multifractal spectra provide a global description of the singularities of the observed measure μ. By construction, $\mu_i(q,r) \in [0,1]$ for each value of q.

The $f(q)-$spectra among groups showed distinctively different shape and symmetry (Fig.3.8). The curvature and the symmetry of the $f(q)-$spectra provide information on the heterogeneity of a system defined by the diversity of scaling exponents needed to characterize it. In this set of images, for each value of q, linear fits to the plots of numerators versus denominators in the above equations were used to calculate $f(q)$, D_q and $\alpha(q)$. q values in the range from -1 to $+1$ were used to calculate $f(q)$ and $\alpha(q)$ as illustrated in Figure 3.8 (a). q values were increased in steps of 0.2. The linearity of fits was assessed using R^2 values. The parameter $\alpha(q)$ quantifies the degree of regularity in a point q, loosely speaking, the measure of an interval $[q, q+\Delta q]$, in this applications, the number of events occurring in this interval. Heterogeneity can be assessed at $q = 0$ by the magnitude of the differences in the values of D_0 and $\alpha(0)$, or more generally, by the magnitude of changes around D_0 in both the $f(\alpha)$ and α axes as shown in the figures 3.8 and 3.9. Figure 3.8 (b) shows the multifractal spectra of the binary images, which were obtained by plotting $f(\alpha)$ against $\alpha(q)$. The spectrum $f(\alpha)$ captures how "frequently" a value $\alpha(q) = q$ is founded. For the almost sure value $\alpha(0)$ one has $f(\alpha(0)) = D_0$

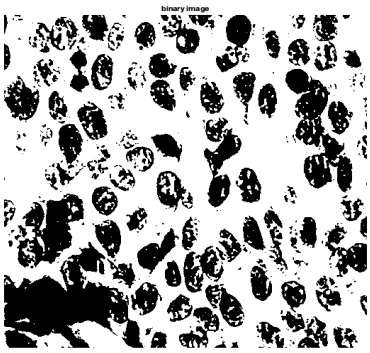


(a) Mastopathic Tissue (A)

(b) Binarizated Mastopathic Image



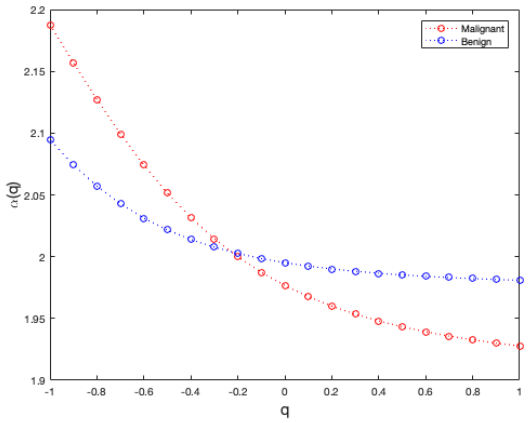
(c) Mammary cancer (DC)



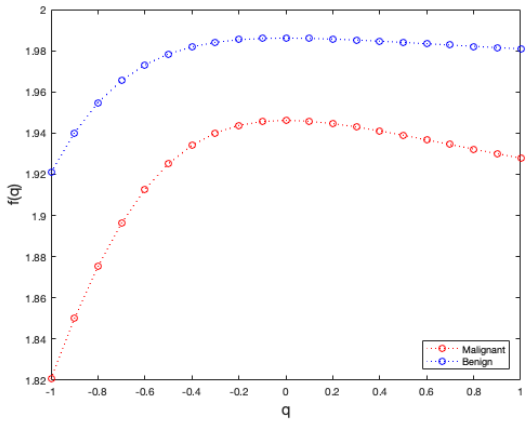
(d) Binarizated Mammary cancer Image

Fig. 3.7 Two original microscopy RGB images of breast tissue from a 400 $\times$ 400 visual field. (a) Mastopathic tissue, in this case adenosis (A) of the breast, which is a benign (non-cancerous) breast condition in which the lobules are enlarged, and there are more glands than usual. (c) Mammary cancer (Ductual carcinoma - DC) is the most common type of breast cancer. *Ductal* means that the cancer began in the milk ducts and *Carcinoma* refers to any cancer that begins in the skin or other tissues that cover internal organs. Finally, (b) and (d). Binary images of (a) and (b) respectively, these were obtained by manual processing technique developed to enhance the visibility of regions of interest and improve intensity distribution.

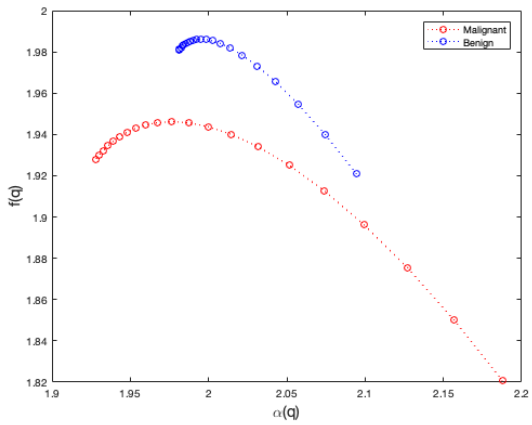
which is necessarily the maximal value of $f(q)$, see Figure 3.8 (c). The $f(q)-$spectra of homogeneous systems with fractal support is reduced to this single point found at $q = 0$, that is, $f[q(0)] = D_0$.



(a) Multifractal spectrum $\alpha(q)$



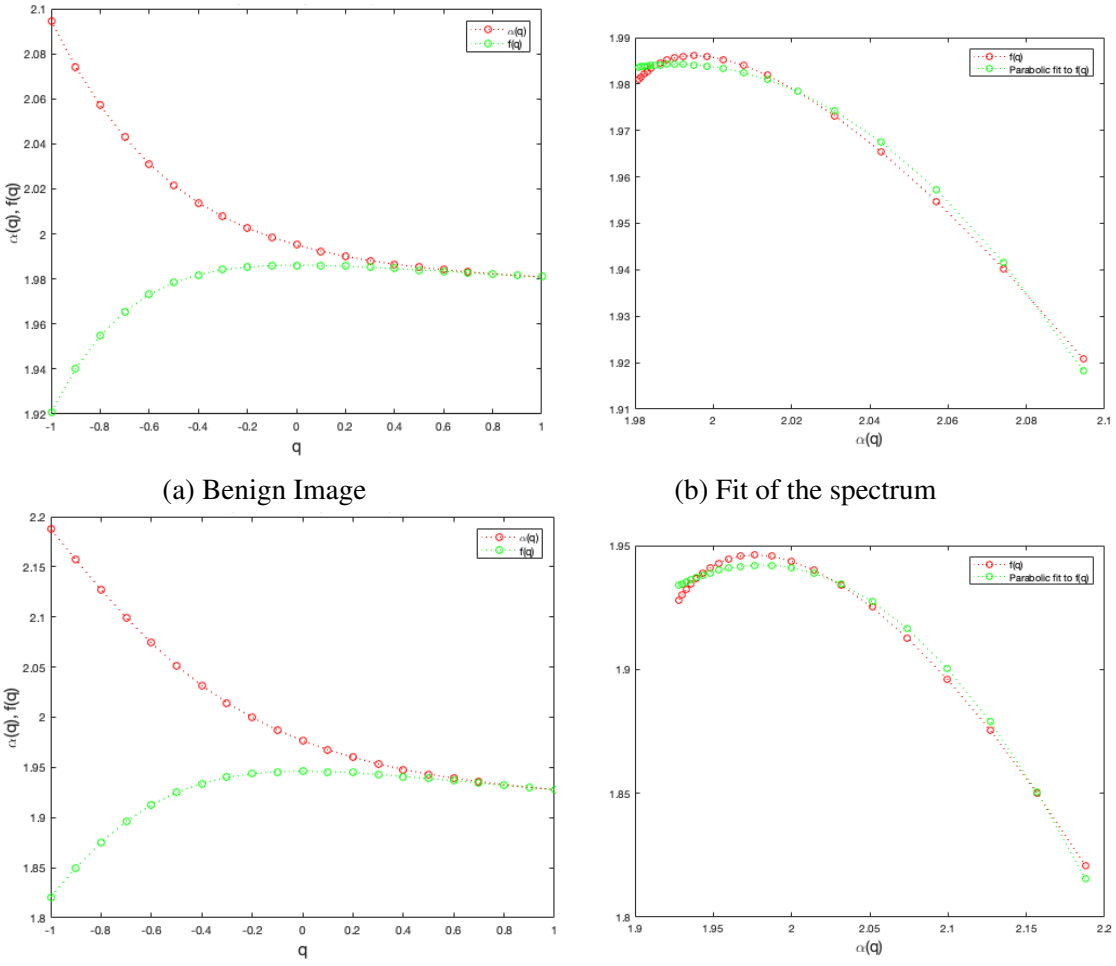
(b) Multifractal spectrum $f(q)$



(c) Multifractal spectra

Fig. 3.8 Multifractal spectrum $\alpha(q)$ (a) and $f(q)$ (b) for each order of the moment q for one image of mammary cancer tissue (red dashes) and one image of masthopatic tissue (blue dashes). (c) Multifractal spectra obtained for the binary images.

In addition to considering the generalized fractal dimensions (D_0, D_1, D_2), the multifractal spectrum can be plotted to visualize the distribution of the space occupied by breast tissue (see Fig. 3.10). Figure 3.8 shows the multifractal spectra corresponding to breast architecture in two classes of tissues. The peak value of the multifractal spectra corresponds to D_0. $f(\alpha)$ values to the left and right of D_0 correspond to positive and



(a) Benign Image

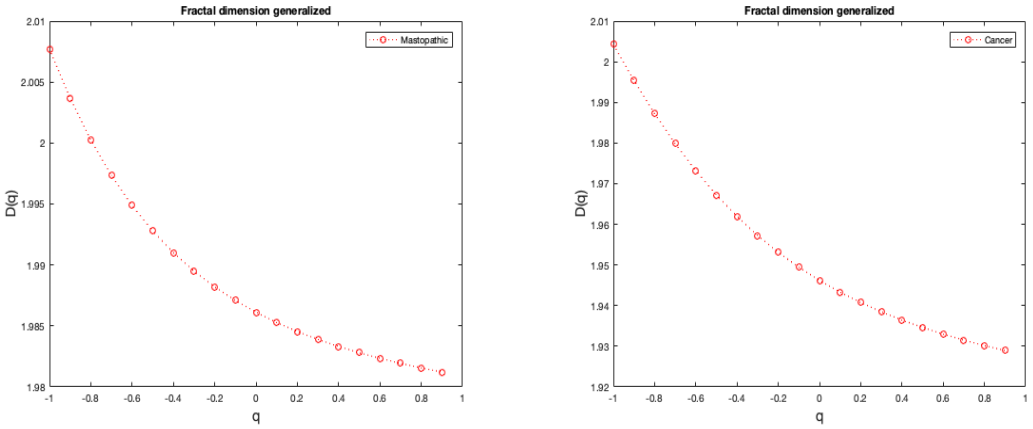
(b) Fit of the spectrum

(c) Malignant Image

(d) Fit of the spectrum

Fig. 3.9 Comparison multifractal spectrum of two benign and malignant images of breast cancer histopathology images (a) and (c) and fit of the singularity multifractal spectrum $\alpha(q)$ versus $f(\alpha(q))$ to a parabola (b) and (d).

negative values of q, respectively. Since $\alpha(q)$ values are estimated from the linear fits to the equations described earlier, values of $q > 0$ and $q < 0$ for which $R^2 > 0.95$ were used for plotting and comparison. The shift in the peaks of the spectra from top to bottom (Figure 3.8c) is due to differences in the values of D_0 (Table 3.1), corresponding to the apex of the spectrum. For example, the malignant tissue occupy more space than the mastopathic tissue, and hence their D_0 values were significantly higher.



(a) (b)

Fig. 3.10 Generalized dimensions of the microscopic images of breast tumor tissue images. All monofractals represent a special case of multifractal behavior in which all generalized dimensions are equal.

Fig. 3.10 represents the generalized dimensions obtained from images 3.9. The plot show that the generalized dimensions of the multifractal structures D_0, D_1 and D_2 had different values, demonstrating the multifractal nature of the microscopic images of breast tumor tissue images.

Riedi (1997) gives some features which are common to all spectra of measures:

- The spectrum of a measure touches the internal bisector of the axis.
- Moreover, for any measure $f(\alpha) \leq \alpha$ for all α.
- The spectrum touches the horizontal line through $(0, D_0)$ where $D_0 = -\beta(0)$.
- Moreover, $f(\alpha) \leq D_0$ for all α.

The purpose of this Chapter is to apply a novel procedure for the direct evaluation of the spectrum $f(\alpha)$ (without resorting to the intermediate Legendre transform), which is

Generalized fractal dimensions			
Tissue	D_0	D_1	D_2
Mastopathic Tissue	1.94	1.07	0.96
Mammary cancer	1.90	1.05	0.97

Table 3.1 Measures of the breast tissue obtained on one image of mammary cancer tissue and one image of masthopatic tissue.

mathematically precise and can be readily applied to the analysis of real experimental data like mammographic images where the underlying dynamics are unknown. This research aims to demonstrate the tissue of the breast as a multifractal object, a fact which will be proven through generalized dimensions D_q and singularity spectrum $f(\alpha)$. Furthermore, it is expected that the singularity spectrum has the capacity to evidence the disorders in the breast architecture with diseases.

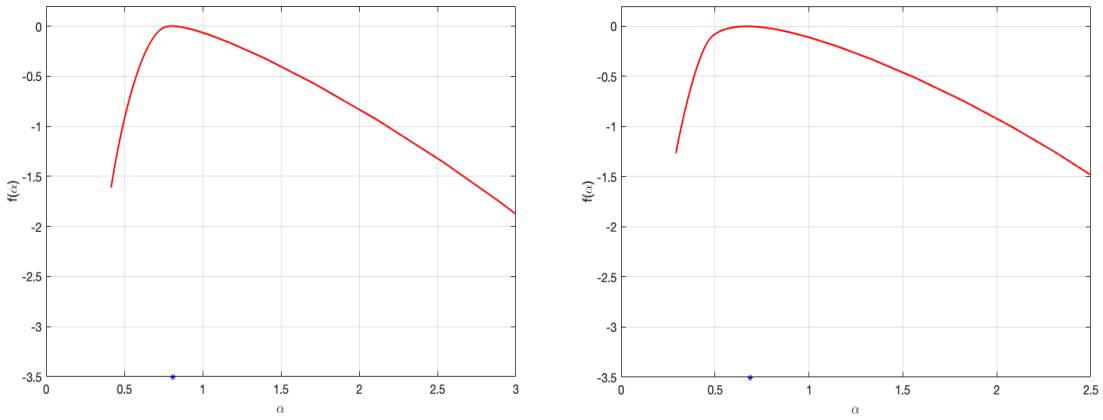
3.5.3 Multifractal analysis of mammography: a wavelet based approach

In this section, we apply the wavelet-based multifractal spectra to two digital mammogram images (shown in Figure 3.9) of size 512×512 representing the mastopathic and cancerous tissues. We refer to the paper of Hermann et al. (2015) for a detailed description of the images.

First, we perform the 2D discrete complex wavelet transform for each image of size 512×512 by using complex Daubechies 6-tap filter (see, Jeon et al. (2014) and Hermann et al. (2015)), then we evaluated the wavelet multifractal spectra by extending the Eqs. 3.19 and 3.22 to 2D.

Figure 3.11 compare the multifractal spectrum of the mastopathic tissue with the cancerous mammogram image. Although they seems having a similar behavior it is evident that the Hurst exponents representing the local regularity are different for the two images.

Here, the different fractality is also confirmed by the calculation of the multifractal descriptors shown in Table 3.2. The mastopathic tissue seems less regular than the cancerous one (the regularity is represented by the SM or Hurst exponent) and the range (or broadness) of the local Hölder index is smaller than the concerous tissue.



(a) (b)

Fig. 3.11 Wavelet multifractal spectrum for the mastopathic tissue (a) and cancerous tissue (b).

Descriptors	H	L1	L2	R1	R2	B
Mastopathic	0.80	3.0	1.32	-0.70	-0.48	0.57
cancer	0.67	3.70	0.90	-0.70	-0.42	0.70

Table 3.2 Wavelet multifractal descriptors.

We think that the multifractal spectrum and its descriptors could be used in classification algorithms for discriminating the mastopathic from the cancerous tissue. This could provide an automatic tool to support medical decisions.

Concluding remarks

1. Percolation is a mathematical concept used to model random processes in a disordered system. In a more technical sense, invasion percolation is an algorithm that models the expansion of a network throughout a medium with randomly distributed heterogeneities in strength. More precisely, the percolation theory describes the permeability of a substance diffusing at random through porous materials and it is also the simplest model to display a state transition phenomenon. While percolation cannot be solved exactly for intermediate dimensions, the model enables the reader to become familiar with important concepts such as fractals, scaling, and renormalisation group theory in a very intuitive way. The percolation theory is independent of the local geometry and dependent on the average of its geometric properties, and the size of the system. Percolation models provide information and predictions on the system's integrity and connectivity which can then be related to the physical or biological tissue properties and structure.

   We showed that invasion percolation can be used both to describe the irregular vascular architecture in tumors. The insights provided by random fractal image analysis and the percolation model provide a general idea about the growth of cancerous tissue in the breast. Irregular structure of cancerous cells can be interpreted as fractals, random carpet (RC) is a flexible parametric model. The fractal analysis is an important tool for the diagnosis of breast cancer since it can quantify the irregularity and complexity of objects. Therein, we realized that deterministic fractals are too artificial models for tissue growth. Here, we used the Percolation model for describing properties related to the tissue's connectivity, too. Both models can be used to describe the irregular architecture in tumors. Therefore, we shall concentrate on models of random carpets and percolation model as a tool to discriminate mammopathic tissues of cancerous tissues in diagnostic and prognostic processes.

2. We sought to differentiate breast tissue in a region of interest (ROI), using texture and fractal analysis from histopatological images. Indeed, this analysis was able to detect differences between cancerous and normal tissue. The fractal and texture parameters which we have addressed the classifications of abnormalities of the breast tissues have an intuitive meaning, and they are a form of characterizing the gray level distributions of each analyzed ROI. These gray level distributions, in turn, underlie physical properties regarding the imaged tissues. We applied our algorithms on breast cancer histopathological images (Spanhol et al., 2016) and we preprocessed images to reduce training time and increase the performance of the classifiers. We used five ML methods, i.e, Decision Tree (DT), Support Vector Machines (SVM), Random Forest (RF), Bayes Networks (BN) and Neural Networks (NN), were tested; and one supervised learning approach (Binary classification). We obtained satisfying classifications rates for cancer and mastopathic tissues. Sensitivity, specificity, accuracy and AUC for comparison of models adjusted and classifiers were showed.

3. We proposed a computarized method for the automated classification of breast lesions in a region of interes (ROI) using some statistical parameters such as the fractal indices, succolarity, lacunarity and classical texture features to analyze breast tissue. Additionally, we implemented advanced machine methods to classify mammography images. The methodology provided an excellent perfomance in classifying the normal versus anormal mammographies reaching a 100% of correct classification. Althouth the results are a little bit worse for classifying abnormalities (benign versus malignant), the proposed methodology based on the combined use of new fractal indicators showed a significant improvement respect the results obtained in the literature on the same data base. We showed the importance of combining fractal and lacunarity features with classical measures of the GLMC in the classification of mammographic images.

4. We had some limitations because conventional statistical learning approaches. In the future, we plan to use convolutional neural network (CNN) based in deep learning algorithms combined with handcrafted image future as the used in this work. Using powerful and automatic approaches may help doctors to have more reliable results with detection and characterization of lesions of the breast tissues.

5. We applied two novel procedure for the direct evaluation of the spectrum $f(\alpha)$ (without resorting to the intermediate Legendre transform), which is mathematically

precise and can be readily applied to the analysis of real experimental data like mammographic images or histopatological images (Suckling et al., 1994; Spanhol et al., 2016), where the underlying dynamics are unknown. Here, we demonstrated the tissue of the breast as a multifractal object, a fact which was proven through generalized dimensions D_q and multifractal descriptors of wavelet based approach. We demostrated that the singularity spectrum has the capacity to evidence the disorders in the breast architecture with diseases.

We used multifractal spectrum and its descriptors in classification algorithms for discriminating the mastopathic from the cancerous tissue. We detected significant differences in the underlying distribution between the two groups. This provided an automatic tool to support medical decisions.

References

Abramovich, F. (2000). Wavelet analysis and its statistical applications. *Journal of the Royal Statistical Society. Series D (The Statistician) 49*(1), 1–29.

Achter, J. D. and C. T. Webb (2006). Pair statistics clarify percolation properties of spatially explicit simulations. *Theoretical population biology 69*(2), 155–164.

Addison, P. S. (1997). *Fractals and chaos: an illustrated course*. CRC Press.

Adler, R. J. (1981). *The geometry of random fields*, Volume 62. Siam.

Albertson, D. G. and D. Pinkel (2003). Genomic microarrays in human genetic disease and cancer. *Human molecular genetics 12*(suppl_2), R145–R152.

Allain, C. and M. Cloitre (1991, Sep). Characterizing the lacunarity of random and deterministic fractal sets. *Phys. Rev. A 44*, 3552–3558.

Arnol'd, V. I. (2013). *Mathematical methods of classical mechanics*, Volume 60. Springer Science & Business Media.

B. Mandelbrot, B. (1983, 03). The fractal geometry of nature.

Barnsley, M. F. and S. Demko (1985). Iterated function systems and the global construction of fractals. *Proc. R. Soc. Lond. A 399*(1817), 243–275.

Breiman, L. (2001). Random forests. *Machine learning 45*(1), 5–32.

Breiman, L. (2017). *Classification and regression trees*. Routledge.

Breiman, L., A. Cutler, A. Liaw, and M. Wiener (2011). Package randomforest. *Software available at: http://stat-www. berkeley. edu/users/breiman/RandomForests*.

Carter, P. H., R. Cawley, and R. D. Mauldin (1988). Mathematics of dimension measurement for graphs of functions. *Proc. Symb. Fractal Aspects of Materials, Disordered Systems*, 183–186.

Chayes, J., L. Chayes, and R. Durrett (1988). Connectivity properties of mandelbrot's percolation process. *Probability theory and related fields 77*(3), 307–324.

Chhabra, A. and R. V. Jensen (1989). Direct determination of the f (α) singularity spectrum. *Physical Review Letters 62*(12), 1327.

Cojocaru, J. I. R., D. Popescu, and I. E. Nicolae (2013). Texture classification based on succolarity. In *Telecommunications Forum (TELFOR), 2013 21st*, pp. 498–501. IEEE.

Cortez, P. (2007). Rminer: Data mining with neural networks and support vector machines using r. *Introduction to Advanced Scientific Softwares and Toolboxes*.

Cortez, P. (2010, July). Data Mining with Neural Networks and Support Vector Machines using the R/rminer Tool. In P. Perner (Ed.), *Advances in Data Mining – Applications and Theoretical Aspects, 10th Industrial Conference on Data Mining*, Berlin, Germany, pp. 572–583. LNAI 6171, Springer.

Cortez, P. (2016). *rminer: Data Mining Classification and Regression Methods*. R package version 1.4.2.

Costa, A. F. (2013). Hausdorff (box-counting) fractal dimension.

Costa, E. V. L. and R. A. Nogueira (2015). Fractal, multifractal and lacunarity analysis applied in retinal regions of diabetic patients with and without nonproliferative diabetic retinopathy. *Fractal Geom Nonlinear Anal Med Biol 1*(3), 112–119.

Courant, R. and D. Hilbert (1954). Methods of mathematical physics, vol. i. *Physics Today 7*(5), 17–17.

Daubechies, I. (1992). *Ten lectures on wavelets*. CBMS-NSF Regional Conference Series in Applied Mathematics. Philadelphia, PA: Society for Industrial and Applied Mathematics.

Davies, S. and P. Hall (1999). Fractal analysis of surface roughness by using spatial data. *Journal of the Royal Statistical Society. Series B 61*(3–37).

De Melo, R. and A. Conci (2008). Succolarity: Defining a method to calculate this fractal measure. In *Systems, Signals and Image Processing, 2008. IWSSIP 2008. 15th International Conference on*, pp. 291–294. IEEE.

De Melo, R. H. and A. Conci (2013). How succolarity could be used as another fractal measure in image analysis. *Telecommunication Systems 52*(3), 1643–1655.

Dekking, F. M. and R. W. J. Meester (1990). On the structure of mandelbrot's percolation process and other random cantor sets. *Journal of Statistical Physics 58*(5-6), 1109–1126.

Derado, G., F. D. Bowman, R. Patel, M. Newell, and B. Vidakovic (2007). *Wavelet image interpolation (WII): A wavelet-based approach to enhancement of digital mammography images*, Volume 4463 of *Lecture Notes in Bioinformatics*, pp. 203–214.

Derado, G., K. Lee, O. Nicolis, F. D. Bowman, M. Newell, F. F. Rugger, and B. Vidakovic (2008). *Wavelet-based 3-D multifractal spectrum with applications in breast MRI images*, Volume 4983 of *Lecture Notes in Bioinformatics*, pp. 281–292.

Dhungel, N., G. Carneiro, and A. P. Bradley (2015). Automated mass detection in mammograms using cascaded deep learning and random forests. In *Digital Image Computing: Techniques and Applications (DICTA), 2015 International Conference on*, pp. 1–8. IEEE.

Dobrescu, R., L. Ichim, and D. Crisan (2013). Diagnosis of breast cancer from mammograms by using fractal measures. *International Journal of Medical Imaging 1*(2), 32–38.

Dobrescu, R. and D. Popescu (2011). Image processing applications based on texture and fractal analysis. In *Applied Signal and Image Processing: Multidisciplinary Advancements*, pp. 226–250. IGI Global.

Dubuc, B., J. Quiniou, C. Roques-Carmes, C. Tricot, and S. Zucker (1989). Evaluating the fractal dimension of profiles. *Physical Review A 39*(3), 1500.

Edgar, G. (2007). *Measure, topology, and fractal geometry*. Springer Science & Business Media.

El-Naqa, I., Y. Yang, M. Wernick, N. Galatsanos, and R. Nishikawa (2002). A support vector machine approach for detection of microcalcifications. *IEEE Transactions on medical imaging 21*(12), 1552–1563.

Falconer, K. (2004). *Fractal geometry: mathematical foundations and applications*. John Wiley & Sons.

Falconer, K. J. (1986). *The geometry of fractal sets*, Volume 85. Cambridge university press.

Falconer, K. J. (1997). Techniques in fractal geometry.

Falconer, K. J. and G. Grimmett (1992). On the geometry of random cantor sets and fractal percolation. *Journal of Theoretical Probability 5*(3), 465–485.

Feder, J. (2013). *Fractals*. Springer Science & Business Media.

Ferraris, L., V. Gabellani, V. Parodi, N. Rebora, J. von Hardenberg, and A. Provenzale (2003). Revisiting multifractality in rainfall fields. *Journal of Hydrometeorology 4*(544-551).

Friedman, J., T. Hastie, and R. Tibshirani (2001). *The elements of statistical learning*, Volume 1. Springer series in statistics New York, NY, USA:.

Gallego-Posada, J., D. Montoya-Zapata, and O. Quintero-Montoya (2016). Detection and diagnosis of breast tumors using deep convolutional neural networks.

Garsia, A. M., E. Rodemich, H. Rumsey, and M. Rosenblatt (1970). A real variable lemma and the continuity of paths of some gaussian processes. *Indiana University Mathematics Journal 20*(6), 565–578.

Gefen, Y., Y. Meir, B. B. Mandelbrot, and A. Aharony (1983, Jan). Geometric implementation of hypercubic lattices with noninteger dimensionality by use of low lacunarity fractal lattices. *Phys. Rev. Lett. 50*, 145–148.

Ghazal, G. A. and H. Neudecker (2000). On second-order and fourth-order moments of jointly distributed random matrices: a survey. *Linear Algebra and its Applications 321*(1-3), 61–93.

Gorshkov, V., V. Privman, and S. Libert (2016). Lattice percolation approach to 3d modeling of tissue aging. *Physica A: Statistical Mechanics and its Applications 462*, 207–216.

Gould, D. J., T. J. Vadakkan, R. A. Poché, and M. E. Dickinson (2011). Multifractal and lacunarity analysis of microvascular morphology and remodeling. *Microcirculation 18 2*, 136–51.

Grassberger, P. and I. Procaccia (1983). Measuring the strangeness of strange attractors. *Physica D: Nonlinear Phenomena 9*(1-2), 189–208.

Haralick, R. M. (1979). Statistical and structural approaches to texture. *Proceedings of the IEEE 67*(5), 786–804.

Heneghan, C., S. Lown, and M. Teich (1996). Two dimensional fractional Brownian motion: Wavelet analysis and synthesis. *Image analysis and interpretation, proceedings of the IEEE Southwest Symposium*, 213–217.

Hermann, P., J. Kiselak, and M. Stehlik (2017). *FractalParameterEstimation: Simulation and Parameter Estimation of Randomized Sierpinski Carpets using the p-p-p-q-Model*. R package version 1.1.1.

Hermann, P., T. Mrkvička, T. Mattfeldt, M. Minárová, K. Helisová, O. Nicolis, F. Wartner, and M. Stehlík (2015). Fractal and stochastic geometry inference for breast cancer: a case study with random fractal models and quermass-interaction process. *Statistics in medicine 34*(18), 2636–2661.

Ho, T. K. (1995). Random decision forests. In *Document analysis and recognition, 1995., proceedings of the third international conference on*, Volume 1, pp. 278–282. IEEE.

Hu, M.-G., J.-F. Wang, and Y. Ge (2009). Super-resolution reconstruction of remote sensing images using multifractal analysis. *Sensors 9*(11), 8669–8683.

Hurst, H. E. (1951). Long-term storage capacity of reservoirs. *Trans. Amer. Soc. Civil Eng. 116*, 770–799.

Jeon, S., O. Nicolis, and B. Vidakovic (2014). Mammogram diagnostics via 2-d complex wavelet-based self-similarity measures. *The São Paulo Journal of Mathematical Sciences 8*(2), 265–284.

Jin, X., S. Ong, et al. (1995). A practical method for estimating fractal dimension. *Pattern Recognition Letters 16*(5), 457–464.

Karssemeijer, N. (1992). Stochastic model for automated detection of calcifications in digital mammograms. *Image and vision computing 10*(6), 369–375.

Kestener, P., J. Lina, P. Saint-Jean, and A. Arneodo (2001). Wavelet-based multifractal formalism to assist in diagnosis in digitized mammograms. *Image analysis and stereology 20*(3), 169–175.

Khemis, K., S. A. Lazzouni, M. Messadi, S. Loudjedi, and A. Bessaid (2016). New algorithm for fractal dimension estimation based on texture measurements: Application on breast tissue characterization. *International Journal of Image, Graphics and Signal Processing 8*(4), 9.

Kohavi, R. et al. (1995). A study of cross-validation and bootstrap for accuracy estimation and model selection. In *Ijcai*, Volume 14, pp. 1137–1145. Montreal, Canada.

Kolmogorov, A. N. (1940). Wienersche spiralen und einige andere interessante kurven in hilbertscen raum, cr (doklady). *Acad. Sci. URSS (NS) 26*, 115–118.

Kourou, K., T. P. Exarchos, K. P. Exarchos, M. V. Karamouzis, and D. I. Fotiadis (2015). Machine learning applications in cancer prognosis and prediction. *Computational and structural biotechnology journal 13*, 8–17.

Lauby-Secretan, B., C. Scoccianti, D. Loomis, L. Benbrahim-Tallaa, V. Bouvard, F. Bianchini, and K. Straif (2015). Breast-cancer screening?viewpoint of the iarc working group. *New England journal of medicine 372*(24), 2353–2358.

Lawton, W. (1993). Applications of complex valued wavelet transforms to subband decomposition. *IEEE Transactions on Signal Processing 41*(12), 3566–3568.

Lehamel, M. and K. Hammouche. Texture classification using fractal dimension and lacunarity.

Liebovitch, L. S. and T. Toth (1989). A fast algorithm to determine fractal dimensions by box counting. *physics Letters A 141*(8-9), 386–390.

Lina, J.-M., M. Mayrand, et al. (1995). Complex daubechies wavelets. *Applied and Computational Harmonic Analysis 2*(3), 219–229.

Lopes, R. and N. Betrouni (2009). Fractal and multifractal analysis: a review. *Medical image analysis 13*(4), 634–649.

Lovejoy, S. and B. Mandelbrot (1985). Fractal properties of rain, and a fractal model. *Tellus 37*(3), 209–232.

Lovejoy, S. and D. Schertzer (2007). Scale, scaling and multifractals in geophysics: Twenty years on. In Tsonis, A and Elsner, JB (Ed.), *NONLINEAR DYNAMICS IN GEOSCIENCES*, pp. 311–337. Aegean Conf; Amer Meteorol Soc; European Geosci Union: SPRINGER. Conference on Nonlinear Dynamics in Geosciences, Rhodes, GREECE, JUN 12-16, 2006.

Lu, Z. and L. M. Berliner (1999). Markov switching time series models with applications to a daily runoff series. *Water Resources Research 35*(2), 523–534.

Mallat, S. (1997). *A Wavelet Tour of Signal Processing*. London: AP Professional.

Mandelbrot, B. B. (1974). Intermittent turbulence in self-similar cascades: divergence of high moments and dimension of the carrier. *Journal of Fluid Mechanics 62*(2), 331–358.

Mandelbrot, B. B. and J. W. Van Ness (1968). Fractional brownian motions, fractional noises and applications. *SIAM review 10*(4), 422–437.

Martin, J., M. Moskowitz, and J. Milbrath (1979). Breast cancer missed by mammography. *American Journal of Roentgenology 37*(2), 142–162.

Melo, R. (2007). Using fractal characteristics such as fractal dimension. *Lacunarity and Succolarity to Characterize Texture Patterns on Images [Master?s thesis]: Niterói, Brazil, Federal Fluminense University, www2. ic. uff. br/PosGraduacao/Dissertacoes/356. pdf*.

Mohamed, A. A., W. A. Berg, H. Peng, Y. Luo, R. C. Jankowitz, and S. Wu (2018). A deep learning method for classifying mammographic breast density categories. *Medical physics 45*(1), 314–321.

Natoli, C. (2012). Fractals as fixed points of iterated function systems. *University of Chicago*.

N'Diaye, M., C. Degeratu, J.-M. Bouler, and D. Chappard (2013). Biomaterial porosity determined by fractal dimensions, succolarity and lacunarity on microcomputed tomographic images. *Materials Science and Engineering: C 33*(4), 2025–2030.

Netsch, T. and H. Peitgen (1999). Scale-space signatures for the detection of clustered microcalcifications in digital mammograms. *IEEE Transactions on medical imaging 18*(9), 774–786.

Nicolis, O., J. Kisel'ák, F. Porro, and M. Stehlík (2017). Multi-fractal cancer risk assessment. *Stochastic Analysis and Applications 35*(2), 237–256.

Nicolis, O., P. Ramírez-Cobo, and B. Vidakovic (2011). 2-D wavelet-based spectra with applications. *Computational Statistics and Data Analysis 55*, 738–741.

Nicolis, O., P. Ramirez-Cobo, and B. Vidakovic (2011). 2d wavelet-based spectra with applications. *Computational Statistics & Data Analysis 55*(1), 738–751.

Nilsson, E. (2007). Multifractal-based image analysis with applications in medical imaging. *Dep. Comput. Sci. Umea Univ. Umea Sweden*, 33–70.

Nordvall, D. (2006). Multifractals in theory and practice.

Nualart, D. (2006). Fractional brownian motion: stochastic calculus and applications. In *International Congress of Mathematicians*, Volume 3, pp. 1541–1562. Eur. Math. Soc.

Parra, C., K. Iftekharuddin, and D. Rendon (2003). Wavelet based estimation of the fractal dimension in fbm images. In *Neural Engineering, 2003. Conference Proceedings. First International IEEE EMBS Conference on*, pp. 533–536. IEEE.

Pignon, D., P. Parmiter, J. Slack, M. Hands, T. Hall, M. Van Daalen, and J. Shawe-Taylor (1996). Sigmoid neural transfer function realized by percolation. *Optics letters 21*(3), 222–224.

Pitsianis, N., G. Bleris, and P. Argyrakis (1989). Information dimension in fractal structures. *Physical review B 39*(10), 7097.

Plotnick, R. E., R. H. Gardner, W. W. Hargrove, K. Prestegaard, and M. Perlmutter (1996). Lacunarity analysis: a general technique for the analysis of spatial patterns. *Physical review E 53*(5), 5461.

Popescu, D., L. Ichim, et al. (2017). Texture classification algorithm using elements of fractal analysis. In *Control Systems and Computer Science (CSCS), 2017 21st International Conference on*, pp. 435–440. IEEE.

Posadas, A. N., D. Giménez, R. Quiroz, and R. Protz (2003). Multifractal characterization of soil pore systems. *Soil Science Society of America Journal 67*(5), 1361–1369.

Privman, V., V. Gorshkov, and S. Libert (2016). Lattice percolation approach to numerical modelling of tissue aging. *International Journal of Parallel, Emergent and Distributed Systems 31*(1), 1–19.

R Core Team (2017). *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing.

Radii, R. and A. Politi (1985). Statistical description of chaotic attractors: the dimension function. *Journal of Statistical Physics 40*(5-6), 725–750.

Ramirez-Cobo, P., K. S. Lee, A. Molini, A. Porporato, G. Katul, and B. Vidakovic (2011). A wavelet-based spectral method for extracting self-similarity measures in time-varying two-dimensional rainfall maps. *Journal of Time Series Analysis 32*(4), 351–363.

Ramírez-Cobo, P. and B. Vidakovic (2013). A 2d wavelet-based multiscale approach with applications to the analysis of digital mammograms. *Computational Statistics & Data Analysis 58*, 71–81.

Rani, M. and S. Aggarwal (2013). Fractal texture: a survey. *Advances in Computational Research 5*(1), 149.

Reljin, I. and B. Reljin (2002). Fractal geometry and multifractals in analyzing and processing medical data and images. *Archive of Oncology 10*(4), 283–293.

Rényi, A. (1955). On a new axiomatic theory of probability. *Acta Mathematica Hungarica 6*(3-4), 285–335.

Rezai-Rad, G. and S. Jamarani (2005). Detecting microcalcification clusters in digital mammograms using combination of wavelet and neural network. In *Computer Graphics, Imaging and Vision: New Trends, 2005. International Conference on*, pp. 197–201. IEEE.

Riedi, R. (1999). Multifractal processes. In M. S. T. Paul Doukhan, George Oppenheim (Ed.), *Theory and Applications of Long-range Dependence*.

Riedi, R., M. Crouse, V. Ribeiro, and R. Baraniuk (1999). A multifractal wavelet model with application to network traffic. *IEEE Transactions on Information Theory 45*, 992–1018.

Riedi, R. H. (1997). An introduction to multifractals. In *Rice University ECE Technical Report*.

Russell, D. A., J. D. Hanson, and E. Ott (1980). Dimension of strange attractors. *Physical Review Letters 45*(14), 1175.

Salat, H., R. Murcio, and E. Arcaute (2017). Multifractal methodology. *Physica A: Statistical Mechanics and its Applications 473*, 467–487.

Sarkar, N. and B. B. Chaudhuri (1992). An efficient approach to estimate fractal dimension of textural images. *Pattern recognition 25*(9), 1035–1041.

Shahzad, F., S. Mahmood, and R. Gloaguen (2010). Nonlinear analysis of drainage systems to examine surface deformation: an example from potwar plateau (northern pakistan). *Nonlinear Processes in Geophysics 17*(2), 137–147.

Spanhol, F. A., L. S. Oliveira, C. Petitjean, and L. Heutte (2016). A dataset for breast cancer histopathological image classification. *IEEE Transactions on Biomedical Engineering 63*(7), 1455–1462.

Stanley, H. E. (2000). Monofractal and multifractal approaches to complex biomedical signals. *AIP Conference Proceedings 502*(1), 133–145.

Steeb, W.-H. (2014a). *The nonlinear workbook: Chaos, fractals, cellular automata, genetic algorithms, gene expression programming, support vector machine, wavelets, hidden Markov models, fuzzy logic with C++*. World Scientific Publishing Company.

Steeb, W.-H. (2014b). *The nonlinear workbook: Chaos, fractals, cellular automata, genetic algorithms, gene expression programming, support vector machine, wavelets, hidden Markov models, fuzzy logic with C++*. World Scientific Publishing Company.

Stehlık, M., P. Hermann, and O. Nicolis (2016). Fractal based cancer modelling. *REVSTAT–Statistical Journal 14*(2), 139–155.

Stehlík, M., J. Kiselak, and D. L. Londoño (2018). Percolation based tissue modeling. In *AIP Conference Proceedings*, Volume 1, 2025, pp. 020005. AIP Publishing.

Stehlík, M., T. Mrkvicka, J. Filus, and L. Filus (2012). Recent developments on testing in cancer risk: a fractal and stochastic geometry. *Journal of Reliability and Statistical Studies 5*, 83–95.

Strang, G. and T. Nguyen (1996). *Wavelets and filter banks*. SIAM.

Suckling, J., J. Parker, D. Dance, S. Astley, I. Hutt, C. Boggis, I. Ricketts, E. Stamatakis, N. Cerneaz, S. Kok, et al. (1994). The mammographic image analysis society digital mammogram database. In *Exerpta Medica. International Congress Series*, Volume 1069, pp. 375–378.

Swishchuk, A. and S. Islam (2016). *Random dynamical systems in finance*. Chapman and Hall/CRC.

Team, R. et al. (2015). Rstudio: integrated development for r. *RStudio, Inc., Boston, MA URL http://www. rstudio. com 42*.

Theiler, J. (1990). Estimating fractal dimension. *JOSA A 7*(6), 1055–1073.

Tolle, C., T. Mcjunkin, and D. J. Gorsich (2008, 03). An efficient implementation of the gliding box lacunarity algorithm. *237*, 306–315.

Vadakkan, T. (2009). Multifractal spectrum of a binari image. matlab central. *MATLAB Central*.

Veitch, D. and P. Abry (1999). A wavelet-based joint estimator of the parameters of long-range dependence. *IEEE Transactions on Information Theory 45*(3), 878–897.

Vicsek, T. (1993). The fractal nature of common patterns. In *Growth Patterns in Physical Sciences and Biology*, pp. 29–36. Springer.

Wang, J., X. Yang, H. Cai, W. Tan, C. Jin, and L. Li (2016). Discrimination of breast cancer with microcalcifications on mammography by deep learning. *Scientific reports 6*, 27327.

Wang, T. and N. Karayiannis (1998). Detection of microcalcifications in digital mammograms using wavelets. *IEEE Transactions on on medical imaging 17*(4), 498–509.

Xiao, Y. (2013). Recent developments on fractal properties of gaussian random fields. In *Further developments in fractals and related fields*, pp. 255–288. Springer.

Yassin, N. I., S. Omran, E. M. El Houby, and H. Allam (2017). Machine learning techniques for breast cancer computer aided diagnosis using different image modalities: a systematic review. *Computer methods and programs in biomedicine*.

Zhang, X.-P., M. D. Desai, and Y.-N. Peng (1999). Orthogonal complex filter banks and wavelets: some properties and design. *IEEE Transactions on Signal Processing 47*(4), 1039–1048.

Zhou, S. K., H. Greenspan, and D. Shen (2017). *Deep learning for medical image analysis*. Academic Press.

Appendix A

A.0.1 Generating random SC

We introduce the following codes for generating random SC in the R package FractalParameterEstimation (Hermann et al., 2017).

```
GSC<-function(p,N, sierp=TRUE)
{
if (N == 1)
return(simMatrix(p, sierp))
else
return(kronecker(GSC(p,N-1,sierp), simMatrix(p, sierp)))
}

GSC_seq<-function(p, sierp=TRUE)
{
N <- length(p)
if (N == 1)
return(simMatrix(p[1], sierp))
else
return(kronecker(GSC_seq(p[1:(N-1)]), simMatrix(p[N], sierp)))
}
```

where procedure **kronecker** is included in base package Matrix and X is 3×3 generating matrix of the fractal, in our case it is given by following procedure

```
simMatrix<-function(p, sierp = TRUE)
{
A<-matrix(rbinom(9,1,p), ncol=3,nrow=3)
```

```
if (sierp) {
A[2,2]<-0
}
return(A)
}
```

Appendix B

B.0.1 Wavelet analysis

We used wavelet-based techniques to analyze the complex structure of the breast tissue in mammogram and histopatological images. These techniques have been adopted to study multi-scale processes, because wavelet theory is based onscale-wise decomposition. The construction of the complex wavelet basis associated with multiresolution analysis follows the usual approach proposed by Mallat (1997) and Daubechies (1992). Details on the construction and properties of complex wavelets can be found in Lawton (1993), Lina et al. (1995), Strang and Nguyen (1996) and Zhang et al. (1999). In particular, wavelets have proven effective in extracting statistical properties of a variety of long-range dependence phenomena, including fractals and other scale-invariant processes, in one or more dimensions. Here we focus on the variance of a 2D spatial process, studied using wavelet decomposition. In analogy to real case, the wavelet function $\psi(x)$ is given by

$$\psi(x) = \frac{1}{\sqrt{2}} \sum_k (-1)^k 2h^*_{1-k} \phi(2x-k).$$

where ϕ is the scaling function, h is the low pass filter and the h^* indicates its complex conjugate. The representation of wavelets in 2D can be done through the tensor product of univariate scaling functions and wavelets as follows

$$\begin{aligned}
\phi(x,y) &= \phi(x)\cdot\phi(y) \\
\psi^h(x,y) &= \phi(x)\cdot\psi(y) \\
\psi^v(x,y) &= \psi(x)\cdot\phi(y) \\
\psi^d(x,y) &= \psi(x)\cdot\psi(y)
\end{aligned} \tag{B.1}$$

where symbols h, v, d in (B.1) stand for horizontal, vertical and diagonal directions, respectively, since the atoms capture image features in the corresponding directions. In the complex domain the scaling and wavelet functions have real and imaginary parts, that is,

$$\begin{aligned}
\phi(x,y) &= \xi^s(x,y) + i\zeta^s(x,y) \\
\psi^i(x,y) &= \xi^i(x,y) + i\zeta^i(x,y)
\end{aligned}$$

for $i = h, v, d$.

B.0.2 The complex scale-mixing 2-D wavelet transform

The discrete complex wavelet transform (DCWT) can be considered a complex-valued extension to the standard discrete wavelet transform (DWT) which uses complex-valued filtering (analytic filter) for decomposing the real/complex signals into real and imaginary parts in transform domain. Complex wavelet coefficients can be computed by using the Mallat's algorithm described in Mallat (1997)

$$c_{j-1,l} = \sum_k h^*_{k-2l} c_{j,k}. \tag{B.2}$$

and

$$d_{j-1,l} = \sum_k g^*_{k-2l} c_{j,k} \tag{B.3}$$

Conversely, the reconstruction is given by

$$c_{j,k} = \sum_l c_{j-1,l} h_{k-2l} + \sum_l d_{j-1,l} g_{k-2l}. \tag{B.4}$$

The complex wavelet atoms can be defined as,

$$\phi_{(j_1,j_2),\mathbf{k}}(\mathbf{x}) = 2^{(j_1+j_2)/2}\ \phi(2^{j_1}x-k_1,2^{j_2}y-k_2) \tag{B.5}$$

$$\psi^i_{(j_1,j_2),\mathbf{k}}(\mathbf{x}) = 2^{(j_1+j_2)/2}\ \psi^i(2^{j_1}x-k_1,2^{j_2}y-k_2), \tag{B.6}$$

where i is one of h, v, or d, and $(j_1,j_2)\in\mathbb{Z}^2$. Then, any function f in $\mathscr{L}_2(\mathbb{R}^2)$ can be represented as

$$\begin{aligned} f(\mathbf{x}) &= \sum_{\mathbf{k}} c_{(J_0,J_0),\mathbf{k}}\ \phi_{(J_0,J_0),\mathbf{k}}(\mathbf{x}) \\ &+ \sum_{j>J_0}\sum_{\mathbf{k}} d_{(J_0,j),\mathbf{k}}\ \psi^h_{(J_0,j),\mathbf{k}}(\mathbf{x}) \\ &+ \sum_{j>J_0}\sum_{\mathbf{k}} d_{(j,J_0),\mathbf{k}}\ \psi^v_{(j,J_0),\mathbf{k}}(\mathbf{x}) \\ &+ \sum_{j_1,j_2>J_0}\sum_{\mathbf{k}} d_{(j_1,j_2),\mathbf{k}}\ \psi^d_{(j_1,j_2),\mathbf{k}}(\mathbf{x}), \end{aligned}$$

and a new 2-D wavelet transform, which we call the *scale-mixing wavelet transform* is obtained. Notice that (j_1,j_2) in (B.5) and (B.6) can be indexed as well as (j_1,j_1+s), where $s\in\mathbb{Z}$. The new scale-mixing detail coefficients are defined as,

$$\begin{aligned} d_{(J_0,j),\mathbf{k}} &= 2^{(J_0+j)/2}\int f(\mathbf{x})\ \psi^{*h}(2^{J_0}x-k_1,2^{j}y-k_2)\,dx\,dy, \\ d_{(j,J_0),\mathbf{k}} &= 2^{(j+J_0)/2}\int f(\mathbf{x})\ \psi^{*v}(2^{j}x-k_1,2^{J_0}y-k_2)\,dx\,dy, \\ d_{(j_1,j_2),\mathbf{k}} &= 2^{(j_1+j_2)/2}\int f(\mathbf{x})\ \psi^{*d}(2^{j_1}x-k_1,2^{j_2}y-k_2)\,dx\,dy, \end{aligned} \tag{B.7}$$

where ψ^{*i} is a complex conjugate of ψ^i. Similar to the traditional one- and two-dimensional cases, the complex scale-mixing detail coefficients are linked to the original image (2-D signal) through a matrix equation. Define by W a $2^n\times 2^n$ wavelet matrix composed by the complex scaling and wavelet filter coefficients h_k andg_k as in Vidakovic 1999 pag. 116 where the wavelet filter is given by $g_k=(-1)^k h^*_{1+N-k}$ and N is a shift parameter which affects the location of the wavelet. Suppose that an $2^n\times 2^n$ image (matrix) A is to be transformed into the wavelet domain. If the rows of A are transformed by a one-dimensional transform given by the wavelet matrix W, then the object WA' represents a matrix in which the columns are transformed rows of A. If the same is repeated on the

rows of WA', the result is

$$B = W(WA')' = WAW'. \tag{B.8}$$

Matrix B will be called the scale mixing or (covariance wavelet transform) of matrix A, and will be the basis for defining the scale-mixing spectra. It represents a finite-dimensional implementation of (B.7) for signal $f(\mathbf{x})$ sampled in a form of matrix A.

The scale-mixing 2-D transform is operationally appealing. Constructing appropriate W is computationally fast and, since W is orthogonal, the inverse transform is straightforward,

$$A = W'BW.$$

In the scale-mixing 2-D wavelet transform the balance of the total energy $E = \text{trace}(AA')$ in the image A, is preserved over the scales and mixture of scales, since the orthogonality of W implies,

$$E = \text{trace}(AA') = \text{trace}(BB'),$$

for $B = WAW'$.

B.0.3 The complex scale-mixing wavelet spectra

The scale-mixing spectra are defined in terms of the complex scale-mixing coefficients (B.7) as

$$S(j) = \log_2 \mathbb{E}\left(|d_{(j,j+s),\mathbf{k}}|^2\right), \tag{B.9}$$

where $j \in \mathbb{Z}$, and $s \in \mathbb{Z}$ is fixed. Notice that the case $s = 0$ in (B.9) corresponds to the traditional diagonal 2-D spectra described in Nicolis et al. (2011).

To calibrate the scale-mixing spectra, consider now a 2-D fBm, $B_H(\mathbf{u})$. For such a process, the scale-mixing detail coefficients are given by

$$d_{(j,j+s);\mathbf{k}} = 2^{j+\frac{s}{2}} \int B_H(\mathbf{u}) \psi^* \left(2^j u_1 - k_1, 2^{j+s} u_2 - k_2\right) d\mathbf{u}, \tag{B.10}$$

where ψ^* denotes the complex conjugate of ψ^d, the wavelet atom in the diagonal direction defined in (B.6). These coefficients are random variables with zero mean and

variance

$$\mathbb{E}\left[|d_{(j,j+s);\mathbf{k}}|^2\right] = 2^{2j+s} \int \psi\left(2^j u_1 - k_1, 2^{j+s} u_2 - k_2\right) \times \psi^*\left(2^j v_1 - k_1, 2^{j+s} v_2 - k_2\right) \mathbb{E}\left[B_H(\mathbf{u}) B_H(\mathbf{v})\right] d\mathbf{u}\, d\mathbf{v}, \tag{B.11}$$

As in Veitch and Abry (1999); Heneghan et al. (1996) and Nicolis et al. (2011), it is assumed here that the coefficients within and across the scales are uncorrelated.

From (B.11), it can be seen that

$$\mathbb{E}\left[|d_{(j,j+s);\mathbf{k}}|^2\right] = 2^{-j(2H+2)}\, V_{\psi,s}(H), \tag{B.12}$$

where $V_{\psi,s}(H)$ is an expression depending on ψ, H and s, but not on the scale j,

$$V_{\psi,s}(H) = -\frac{\sigma_H^2}{2} \int\int \psi(\mathbf{p}+\mathbf{q})\; \psi^*(\mathbf{q})\; |\mathbf{p}_s|^{2H} 2^{-s}\, d\mathbf{p}\, d\mathbf{q}, \tag{B.13}$$

A proof of (B.12) is provided in the Appendix. By taking logarithms in (B.12),

$$\log_2 \mathbb{E}\left[|d_{(j,j+s);\mathbf{k}}|^2\right] = -(2H+2)\, j + \log_2 V_{\psi,s}(H), \tag{B.14}$$

for $j \in \mathbb{Z}$, and thus the Hurst exponent can be estimated from the slope of the linear equation (B.14). Finally, the empirical counterpart of (B.14) is a regression defined on

$$\left(j, \log_2 \frac{1}{n} \sum_{j,j+s} \left|d_{(j,j+s),\mathbf{k}}\right|^2\right), \quad j,s \in \mathbb{Z}. \tag{B.15}$$

Instead of the sample mean in (B.15), a more robust location measure could be used, such as the median.