

Theory, Analyses and Predictions of Multifractal Formalism and Multifractal Modelling for Stroke Subtypes' Classification

Yeliz Karaca[1(✉)], Dumitru Baleanu[2,3], Majaz Moonis[1], and Yu-Dong Zhang[4]

[1] University of Massachusetts Medical School, Worcester, MA 01655, USA
yeliz.karaca@ieee.org, Majaz.Moonis@umassmemorial.org
[2] Department of Mathematics, Çankaya University, 1406530 Ankara, Turkey
dumitru@cankaya.edu.tr
[3] Institute of Space Sciences, Magurele, Bucharest, Romania
[4] Department of Informatics,University of Leicester, Leicester LE1 7RH, UK
yudongzhang@ieee.org

Abstract. Fractal and multifractal analysis interplay within complementary methodology is of pivotal importance in ubiquitously natural and man-made systems. Since the brain as a complex system operates on multitude of scales, the characterization of its dynamics through detection of self-similarity and regularity presents certain challenges. One framework to dig into complex dynamics and structure is to use intricate properties of multifractals. Morphological and functional points of view guide the analysis of the central nervous system (CNS). The former focuses on the fractal and self-similar geometry at various levels of analysis ranging from one single cell to complicated networks of cells. The latter point of view is defined by a hierarchical organization where self-similar elements are embedded within one another. Stroke is a CNS disorder that occurs via a complex network of vessels and arteries. Considering this profound complexity, the principal aim of this study is to develop a complementary methodology to enable the detection of subtle details concerning stroke which may easily be overlooked during the regular treatment procedures. In the proposed method of our study, multifractal regularization method has been employed for singularity analysis to extract the hidden patterns in stroke dataset with two different approaches. As the first approach, decision tree, Naïve bayes, kNN and MLP algorithms were applied to the stroke dataset. The second approach is made up of two stages: i) multifractal regularization (kulback normalization) method was applied to the stroke dataset and mFr_stroke dataset was generated. ii) the four algorithms stated above were applied to the mFr_stroke dataset. When we compared the experimental results obtained from the stroke dataset and mFr_stroke dataset based on accuracy (specificity, sensitivity, precision, F1-score and Matthews Correlation Coefficient), it was revealed that mFr_stroke dataset achieved higher accuracy rates. Our novel proposed approach can serve for the understanding and taking under control the transient features of stroke.

© Springer Nature Switzerland AG 2020
O. Gervasi et al. (Eds.): ICCSA 2020, LNCS 12250, pp. 410–425, 2020.
https://doi.org/10.1007/978-3-030-58802-1_30

Notably, the study has revealed the reliability, applicability and high accuracy via the methods proposed. Thus, the integrated method has revealed the significance of fractal patterns and accurate prediction of diseases in diagnostic and other critical-decision making processes in related fields.

Keywords: Multifractal Formalism · Fractional brownian motion · Hurst exponent · Fractal pattern · Self-similar process · Stroke · Prediction algorithms · Naïve Bayes algorithm, kNN algorithm, Multilayer Perceptron Algorithm, Multifractal regularization

1 Introduction

Do fractals exhibit ubiquitous patterns? Fractal is the common term for the complex geometric shapes in mathematics that are self-similar with bizarre fragmented patterns. There are key features of fractals, the first is self-similarity, having an iterative nature and level of irregularity and fragmented dimension [1,2]. Fractals are utilized for measuring peculiar phenomena which are challenging to describe objects inclined to repeat themselves on varying scales or which display self-similarity. As fractal analysis is an essential way of measuring various phenomena, it is applied to multiple fields such as economy [3–5], geology [6], space science [7], materials technology [8], epidemiology[9], signal processing (EEG/ECG) [10], diagnostic imaging [11] and brain structure [12–14]. Taking these features into consideration, fractals are ubiquitous in the world, from natural objects to the universal structures, from smallest items to largest scales, including the brain as the most complex organ. For a thorough understanding of the brain and its dynamics, quantification is required since the brain operates on more than one scale and fractal geometry comes to play to address the related challenges.

Stroke is a major CNS disorder leading to death following cancer and cardiovascular disorders, ranking 5th among death causes in the USA. Stroke occurs within a complex network of vessels and arteries [16]. Characterized by a sudden interruption in the blood supply of the brain, strokes are mostly are caused by an immediate blockage of arteries leading to the brain or bleeding into brain tissue. Strokes can display varying symptoms, so they can be difficult to diagnose. In this study, four subtypes of stroke have been handled, which are Large vessel, Small vessel, Cardioenbolic and Cryptogenic (for further details on the subtypes see [12,16]).

When the studies in which fractal and multifractal methods are used for stroke disease are investigated, it is seen many studies yield efficient outcomes exist. The study by [17] is on acute ischemic stroke, revealing an association between decreased fractal dimension of heart rate variability as well as recurrent ischemic stroke following acute ischemic stroke. Their study highlights the importance of the predictive value of fractal dimension. Another study is by [18]

and the results yield the importance of fractal parameters of plaque for determining its severity to facilitate diagnostic processes. The comprehensive review study of [19] provides a sketch of related studies on fractals, neurodegeneration and stroke.

Considering the studies on fractal and multifractal analyses, there exists a large body of applications in the literature. To start with medicine, the study by [20] is concerned with retinal vascular fractal dimension in bipolar disorder and schizophrenia by box-counting method utilizing automated algorithm. [21] developed a new methodology for automated differential diagnosis of Alzheimer's using EEG signals with an investigation of three measures of fractality. Besides applications of fractals in medicine, there are studies in other fields. To illustrate, the study by [22] proposed a model for predicting fractal dimensions at different heights of mining. The authors' model based on the fractal theory achieved appropriate estimation of the position of high-level boreholes. The study by [23] on surface profile reveals that the estimated fractal dimension increased with the sampling length. Another study on reservoirs [24] utilized fractal geometry to develop new models for complex-structured reservoirs by various mathematical dimension types and to develop instantaneous source function via fractal geometry. Additionally, recent works focus on fractal use in different fields, for example, sign language. The study by [25] is on the fractal analysis of a Czech sign language text on three levels of scaling, considering the sign language in terms of fractal dimensions and Hurst exponents.

Precision in forecasting plays a critical role in medicine for patients' survival and life quality. The study by [26] reveals that computer-aided systems with Artificial Intelligence and advanced signal processing techniques can assist physicians with their analyses and interpretation physiological signals and images effectively. The study by [27] addresses the timely prediction of stroke by Principal component analysis (PCA). Another study by [28] on pneumonia investigates post-stroke pneumonia prediction models with the utilization of more cutting-edge machine learning algorithms. Their predictive model is found out to be feasible for the management of stroke, achieving optimal performance when compared with traditional machine learning methods. The study by [29] provides the evaluation of machine learning outcome prediction for stroke. The study results demonstrate that machine learning techniques on extensive datasets were able to predict the functional stroke outcome. Regarding the use of ensemble methods, the study by [30] deals with multiple classification algorithms for stroke data prediction. The experimental results revealed that with the classifier Ensemble, higher prediction accuracy was achieved. The study by [14] provides efficient clustering algorithms' application with 2D multifractal denoising techniques (Bayesian (mBd), Nonlinear (mNold) as well as Pumping (mPumpD)). The results of the study revealed that 2D mBd technique was the most efficient feature descriptor in terms of accuracy for each subtype of stroke. Concerning the mobile application for stroke dataset, the study of [15] designed a mobile and server application for two stroke subtypes. In their proposed model, they

used Multilayer Perceptron Algorithm (MLP), which produced an efficient and informative services framework for stroke determination and management.

Since brain and related diseases present challenges, it is important to generate optimal and reliable methods. Earlier works [17–30] in the literature include classification with machine learning methods; yet, no work exists in the literature in which a multifractal method and comparative analysis methods (decision tree, Naïve bayes, kNN and MLP) have been applied to such an extensive stroke dataset. Both in terms of the dataset studied and the method employed, this study is the first one of its kind for the classification of four stroke subtypes, which are Large vessel, Small vessel, Cardioenbolic and Cryptogenic. Within this framework, the main aim of our study is to develop a balancing methodology for the detection of subtle details of stroke. For this purpose, multifractal regularization was utilized for singularity analysis to extract the hidden patterns in stroke dataset by two different approaches. While the first approach includes the application of decision tree, Naïve bayes, kNN and MLP algorithms to the stroke dataset, the second approach has two stages in itself: firstly, multifractal regularization (kulback normalization) was applied to the stroke dataset and mFr_stroke dataset was generated. Subsequently, the four algorithms stated above were applied to the mFr_stroke dataset.

The rest of the paper is organized as follows. Section 2 deals with the Materials and Methods of the study. Methods of our integrated approach are Multifractal Formalism and Multifractal Analysis, multifractal regularization (Kulback norm) technique in stroke dataset experiments and Prediction Models Using Algorithms: Decision Tree, Naïve Bayes, kNN and MLP. As the subsequent section, experimental results and discussion are provided along with the explanations. As the final section, namely Sect. 4, conclusion is presented.

2 Materials and Methods

2.1 Patient Details

For this study, 1926 individuals were kept under observation by Massachusetts Medical School, University of Worcester, Massachusetts, USA. Compared to the strokes in the nondominant hemisphere, the ischemic strokes in the dominant hemisphere bring about more functional deficits based on the assessment by the National Institutes of Health Stroke Scale (NIHSS).

The total number of ischemic stroke patients, included in our experiments, is 1926 patients (with males [labelled by (1)] and with females [labelled by (0)]). The age of the ischemic stroke patients range from 0 to 104, displaying seven ischemic stroke subtypes (Large vessel (481), Small vessel(228), Cardioenbolic(689), Cryptogenic (528)) as dealt with in this study (see Hindawi for age distribution details of stroke patients). In this study, demographic information, medical history, results of laboratory tests, treatments, and medications data, as can be seen in Table 1, pertaining to 1926 stroke subtypes patients. Table 2 provides the main headings of attributes used for the stroke subtypes.

Table 1. The stroke dataset with the features.

Number of stroke subtypes / TOAST	Main heading of attributes	Data size
Large vessel (481) Small vessel (228) Cardioembolic (689) Cryptogenic (528)	Demographic information (Age, gender) Medical history (HTN, hyperlip, DM, H/O stroke/TIA, AtrialFib, CAD, CHF, PAD/carotid disease, tobacco, ETOH) Results of laboratory test (mRS 90 days, hemorrhagic conversion, NIHSS admission, TPA) Treatment and medication data (Statin, antiHTN, antidiabetic, antiplatelet, anticoagulation, CT perfusion, neurointervantion)	1926×23

TOAST: type/etiology of stroke; TIA; ischemic attack: HTN: hypertension; DM: diabetes mellitus; CAD: coronary artery disease; AtrialFib: atrial fibrillation Stroke; CAD: coronary artery disease; CHF: congestive heart failure; PAD/carotid disease: peripheral arte disease; NIHSS 90 days: National Institutes of Health Stroke Scale 90-day mortality; CT perfusion: computer tomography perfusion, ETOH: alcohol; antiHTN: antihypertensive drugs after acute ischemic stroke; NIHSS discharge: National Institutes of Health Stroke Scale; H/O stroke/TIA: history of transient ischemic attack.

The details on the stroke dataset attributes as used in our study as well as the corresponding descriptions can be seen in [14,15].

2.2 Methods

The key aim of our study is to devise and develop a complementary methodology which could enable the detection of subtle details regarding stroke that may easily be disregarded or overlooked during the regular treatment procedures. Considering the profound complexity of brain, such cases are in question particularly in complex and dynamic structures such as in the case of stroke, which is a disease related to brain. In the proposed method of our study, multifractal regularization method has been employed for singularity analysis to extract the hidden patterns in stroke dataset with two different approaches whose steps are specified below:

(i) The first approach is made up of two stages: a) multifractal regularization (kulback normalization) method was applied to the stroke dataset (X) and mFr_stroke dataset ($\hat{X}$ was generated. b) the four algorithms stated in the first approach were applied to the mFr_stroke dataset.
(ii) As the second approach, decision tree, Naïve bayes, kNN and MLP algorithms were applied to the stroke dataset.
(iii) We compared the results obtained from the first and second approach regarding the stroke dataset and mFr_stroke dataset based on accuracy rate (specificity, sensitivity, precision, F1-score and Matthews Correlation Coefficient).

As a result of all these applications, it has been revealed that mFr_stroke dataset achieved higher accuracy rates for the identification of the subtypes

of stroke. Our novel proposed approach can serve for the understanding and taking under control the transient features of stroke while managing the process efficiently.

All the analyses were computed and the figures of the analyses for the study were performed by Matlab [31] and FracLab [32].

Multifractal Formalism and Multifractal Analysis. Common in nature, multifractal systems are a generalisation of a fractal system where a single exponent like the fractal dimension would not prove to be adequate to describe its dynamics. Therefore, continuous spectra of exponents, namely singularity spectrum, would be required [33]. Multifractal analysis, which includes the distortion of datasets extracted from patterns, is employed to analyse datasets. In this way, multifractal spectra will be generated, presenting the way scaling changes over the dataset. Multifractal analysis techniques have various applications, particularly ones regarding prediction of natural phenomena and interpretation of medical images [34,35]. Besides these techniques, multifractal denoising is a regularization technique which places a local restriction on the reconstructed signal. Instead of requiring that the denoised signal pertains to global smoothness class, a regularized signal with prescribed Hölder exponent is to be sought.

Concerning this study, the enhancement or denoising of complex data, namely the stroke dataset, depends on the analysis of the local Hölder regularity. It is herein supposed that data enhancement is analogous to increasing the Hölder regularity at each point [36]. These methods are aligned well to the cases where sort of data is to be recovered are very irregular, for instance, not differentiable with local regularity varying quickly [14].

The most simplified concept of smoothness pertaining to a function is provided via the C^k differentiability. In this context, a bounded function f belonging to $C^1(R^d)$ provided that it has partial derivatives $\partial f/\partial x_i$ everywhere that display continuity and being bounded; while C^k differentiability concerning $k \geq 2$ is defined by the recursion: f belonging to $C^k(R^d)$ [11]. In this case, should it belong to $C^1(R^d)$ and all of its partial derivatives $\partial f/\partial x_i$ belong to $C^{k-1}(R^d)$, the provision of a definition will be for uniform smoothness (in case the regularity exponent k is an integer. Taylor's formula follows the definition of C^k differentiability, conveying that, for any $x_0 \epsilon R^d, C > 0, \delta > 0$ exists and a polynomial P_{x_0} of degree is lower than k (see [37]). This C^k differentiability regarding the consequence is in the accurate form so that a definition of pointwise smoothness could be yielded, which could also be applicable for fractional orders of smoothness. This outcome was converted into a definition ensuing a usual process in mathematics [37].

Definition 2.1. $\alpha \geq 0$, and $x_0 \in R^d$ is a function and $f : R^d \to R$ is $C^\alpha(x_0)$ should there exist $C > 0$, $\delta > 0$ and a polynomial P_{x_0} of degree which is lower than α [37]:

$$if\ |x - x_0| \leq \delta, \text{then}\ |f(x) - P_{x_0}(x)| \leq C|x - x_0|^\alpha \tag{1}$$

The Hölder exponent of f at x_0 is $h_f(x_0) = \sup\{\alpha : fisC^\alpha(x_0)\}$ [37].

Among the most broadly employed notions of pointwise regularity, we see the Hölder regularity (see [37] for the step details).

As a last step, $B_r = \{x : |x - x_0| \leq r\}$. It could also be noted that the definition of pointwise Hölder regularity can be weakened, so in this way it is possible to notice that (1) can be reexpressed like this [37]:

$\|f - P_{x_0}\|_{B_{r,\infty}} \leq Cr^{\alpha}$

Through multifractal analysis, the distribution of pointwise regularities (or singularities) of irregular functions f on R^m can be described in a statistical and geometrical manner. The conventional concept of notion of pointwise regularity, as the most frequently employed one is the Hölder [12] (for further details on proceeding equations and theoretical aspects see the references ([13,37–39]).

Jaffard and Mélot [40,41] have verified the following: let Ω be a domain of R^m while y is located at the boundary $\partial\Omega$ of Ω, then through taking $P = 0$ or $P = 1$, the condition of $T^p_{\alpha/p}(y)$ pertaining to the characteristic function x_Ω shall concur with the weak α accessibility of Ω bilaterally at y. (assuming $\alpha \geq 0$). Ω is defined to be bilaterally weak accessible at y should there be the existence of two constants C as well as $0 < R < 1$, which is denoted as in (2):

$$\begin{aligned} &\forall r \leq R \\ &\min\left\{meas\left(\Omega^c \cap B\left(y,r\right)\right), meas\left(\Omega \cap B\left(y,r\right)\right)\right\} \\ &\leq Cr^{\alpha+m}, \end{aligned} \tag{2}$$

Here, *meas* refers to the Lebesgue measure, which would enable the applicability of performing of a multifractal analysis regarding the fractal boundaries [42,43]. Such an analysis also offers numerous applications in different areas including mechanics, physics or chemistry (biology, medicine and so on) in which a lot of occurrences encompass fractal interfaces [39] for further details on proceeding equations and theoretical aspects see the references ([14,41] and [44–46]).

The multifractal analysis of a particular dataset entails the definition of a function, performing the computation of its multifractal spectrum as well as the classification of every point in line with the corresponding value of $(\beta, f_h(\alpha))$ geometrically. The α value provides local information concerning the pointwise regularity. The value of $f_h(\alpha)$ provides global information. In multifractal denoising, no assumption is made as to the noise structure or how the noise structure interacts with the veri. $s(x, y)$ represents the original signal, and $n(x, y)$ is the noise. $w(x, y)$ denotes the observed signal while $z(x, y)$ is the estimate of the signal, and (x, y) referring to the pixel location [47]. Based on the dataset in our study, a regularized version $z(x, y)$ of $w(x, y)$ fulfilled the following conditions [47].

(1) $z(x, y)$ has proximity with $w(x, y)$ in the multifractal regularization technique sense (Kullback Norm), which comes to mean that $\|z(x, y) - w(x, y)\|_L$, is minimum. Kullback Norm was applied in our study since and distance was minimized by utilising the Kullback norm. One advantage of this is that the calculations become simpler as an analytical minimization replaces the numerical one, which enables further generalizations to be made.

(2) $(\hat{X})$'s local Hölder function is specified.

$\alpha_z = \alpha_w + \delta$, in which $\alpha_z = \alpha_s$ denotes the estimated signal's Hölder regularity, δ refers to a positive function that is user-defined. Thus, the regularity of shall have a larger value than one of the observations everywhere. And, the estimation of the Hölder exponent is done based on a wavelet-based procedure [47]. In the current study, multifractal regularization (Kulback norm) technique was applied to stroke dataset (X). After the application of multifractal regularization to the stroke dataset, mFr_stroke dataset $(\hat{X})$ was generated as a result of regularity-based enhancement $(\hat{X})$ from multifractal regularization (Kullback norm) technique.

2.3 Prediction Models Using Algorithms: Decision Tree, Naïve Bayes, kNN and MLP

Prediction refers to making estimations regarding future relying on previous and current data. Risk and uncertainty are the key elements of prediction and estimation. Estimation refers to the process of finding an estimate in a setting in which the population parameter could be either unstable or uncertain. Accordingly, in our study, four different algorithms [48] (Decision tree algorithm, Naïve bayes algorithm, kNN algorithm and MLP algorithm) were applied for prediction modelling purposes in line with the proposed method for the stroke subtypes' classification. Multilayer perceptron (MLP) Multilayer perceptron (MLP) belongs to a class of feedforward artificial neural network (ANN). An MLP has minimum three layers of nodes, which are an input layer, a hidden layer as well as an output layer. Each node, except for the input, is a neuron which utilises an activation function that is non-linear. Each neuron in the network takes the arithmetic mean of the weight vectors terminated at it. After that, the outcome is transferred to all the neurons in the subsequent layer depending on the activation function used [15,47].

As for the learning technique, MLP employs a supervised learning technique which is referred to as backpropagation for the purposes of training ([47,48] MLP ile ilgili makale). As a superiority compared to a linear perceptron, the non-linear activation and multiple layers of MLP enable the distinguishing the data which are not separable.

The denotations used for the recursion of the weights based on the assumption that sigmoid activation function is used in between the layers are specified in 3. Let for each neuron in the output layer have a neuron output [15,48] and [49].

$$O_k = \frac{1}{1 + e^{-net_K}} \tag{3}$$

Here, O_k signifies the activation value of the output layer (as in 4) [15,48] and [49].

$$net_k = \sum W_{jk} O_j \tag{4}$$

The denotation of the activation values for the hidden layer is specified in 5 and 6.

$$O_j = \frac{1}{1 + e^{-net_j}} \tag{5}$$

$$net_j = \sum_i W_{ij} O_i \tag{6}$$

The recursion of the weights is realized as in 7.

$$W_{jk} = W_{jk} + \Delta W_{jk} \tag{7}$$

Here, ΔW_{jk} refers to the weight recursion value. In the backpropogation algorithm, the error criterion can be used, which is also known as mean square error (see [15,48] and [49] for further details concerning mean square error, dependency of the error on the weights and recursion).

3 Experimental Results and Discussion

The novel approach we proposed in our study serves for the understanding and taking under control the transient features of stroke which is a CNS disorder occurring through a complex network of vessels and arteries. Considering the fact that brain and related diseases pose challenges, generating optimal and reliable methods is of importance. To serve our purpose, have adopted two approaches. The first includes the application of decision tree, Naïve bayes, kNN and MLP algorithms to the stroke dataset. The second approach has two stages: i) multifractal regularization (kulback normalization) was applied to the stroke dataset (X) and mFr_stroke dataset ($\hat{X}$) was generated. ii) the four aforementioned algorithms were applied to the mFr_stroke dataset ($\hat{X}$). The accuracy rates, specificity, sensitivity, precision, F1-score and Matthews Correlation Coefficient were computed. Thus, predictions were performed for the classification of stroke subtypes. Accordingly, Table 2 provides the steps of our proposed integrated approach.

The procedural steps of the proposed approach of our study are specified below in their respective order:

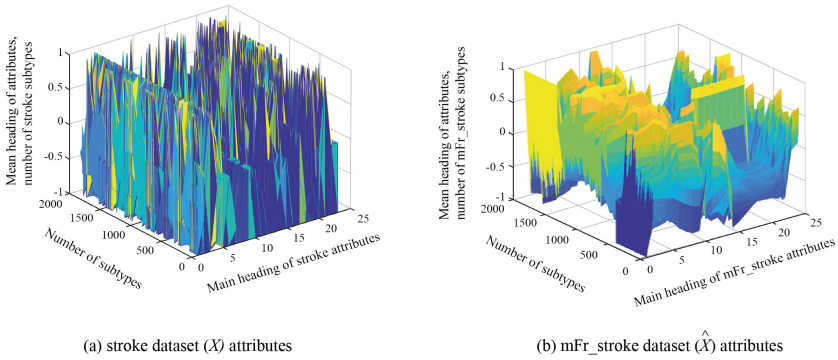
Step 1: Multifractal regularization technique (Kulback normalization) was applied to the stroke dataset (X) **and mFr_stroke dataset** ($\hat{X}$) **was generated: the Display of stroke dataset** (X) **and mFr_stroke dataset** ($\hat{X}$) **on Meshplots.**

In this study, the stroke dataset has a dimension matrix of (1926×23) (see Table 1 for details) and to the stroke dataset (X), multifractal regularization (kulback normalization) was applied as a result of which mFr_stroke dataset ($\hat{X}$) was generated to detect self-similarity and regularity. Figure 1(a) provides the main attributes regarding the stroke dataset. After the multifractal regularization (Kulback normalization) technique was applied to the stroke dataset (X)

Table 2. The steps of the proposed integrated approach.

Step 1: Multifractal regularization (Kulback normalization) was applied to the stroke dataset (X) and mFr_stroke dataset ($\hat{X}$) was generated: the Display of stroke dataset (X) and mFr_stroke dataset ($\hat{X}$) on Meshplots.
Step 2: The Confusion Matrix Computations of Decision Tree, Naïve Bayes, kNN and MLP algorithms for the classification of stroke subtypes on stroke dataset (X) and mFr_stroke dataset ($\hat{X}$).
Step 3: The accuracy rate results based on decision tree, Naïve bayes, kNN and MLP algorithms for the classification of stroke subtypes on stroke dataset (X) and mFr_stroke dataset ($\hat{X}$)

and mFr_stroke dataset ($\hat{X}$) was generated; and Fig. 1(b) presents the multifractal regularity that corresponds to all the attributes in the mFr_stroke dataset through the illustration of meshplots. As Fig. 1(b) depicts, the self-similar and regular patterns have been detected by multifractal regularization, and the theoretical spectrum with increments have been visually illustrated.

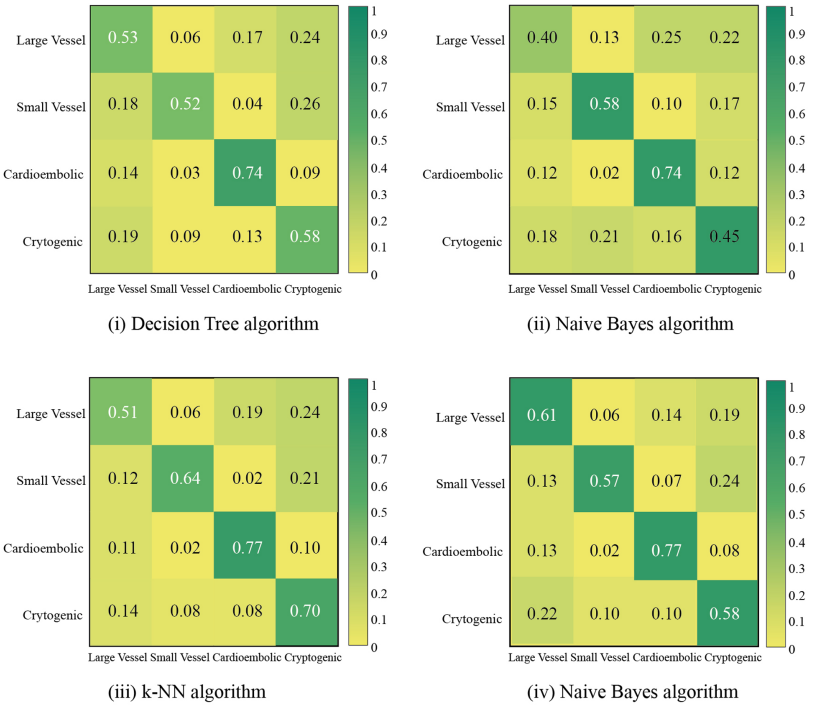


(a) stroke dataset (X) attributes (b) mFr_stroke dataset ($\hat{X}$) attributes

Fig. 1. Display by Meshplot for (a) stroke dataset (X) attributes (b) mFr_stroke dataset ($\hat{X}$) attributes.

Step 2: The Confusion Matrix Computations of Decision Tree, Naïve Bayes, kNN and MLP algorithms for the classification of stroke subtypes on stroke dataset (X) and mFr_stroke dataset ($\hat{X}$).

In our study, confusion matrices were computed as a result of the application of the four algorithms for prediction purposes regarding the classification of stroke subtypes (see Fig. 2 and Fig. 3). Correspondingly, Fig. 2 provides the confusion matrix results for four algorithms and four stroke subtypes for the stroke dataset. Figure 3 presents the confusion matrix results for four algorithms and four stroke subtypes for the mFr_stroke dataset ($\hat{X}$).

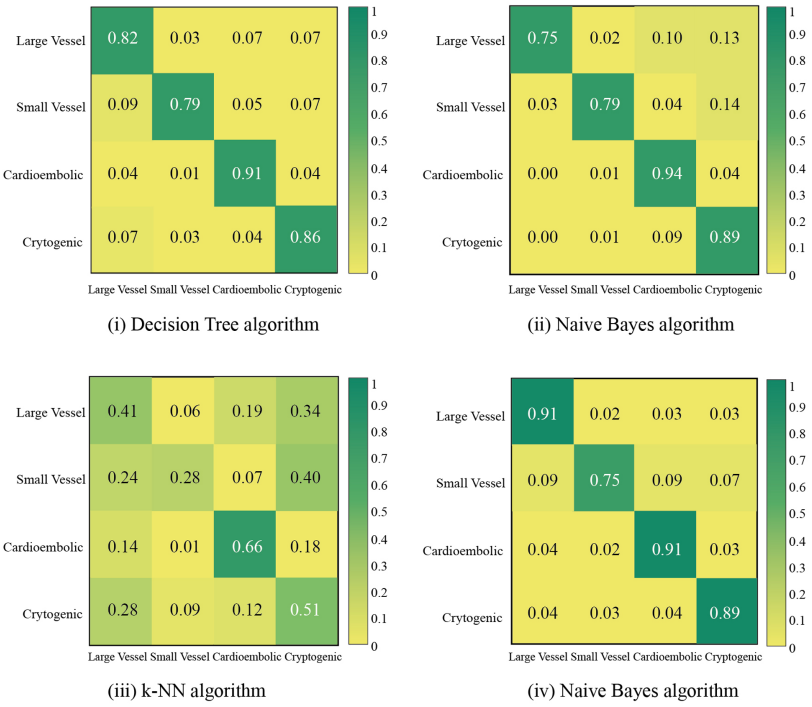


(a) Confusion matrix results for four stroke subtypes in stroke dataset (X) with (i) Decision Tree (ii) Naive Bayes (iii) k-NN (iv) MLP

Fig. 2. Confusion matrix results for four stroke subtypes in stroke dataset (X) and mFr_stroke dataset ($\hat{X}$) using (i) Decision tree (ii) Naïve Bayes (iii) k-NN (iv) MLP.

Table 3. The accuracy rate results based on the application of the algorithms for (a) stroke dataset (X) (b) mFr_stroke dataset ($\hat{X}$).

(a) stroke dataset (X)							
Algorithms	Accuracy rate	Sensitivity	Specificity	Precision	False Positive Rate	F1- score	MCC
Decision Tree	0.616	0.561	0.868	0.593	0.131	0.592	0.461
Naïve Bayes	0.554	0.520	0.849	0.540	0.150	0.524	0.378
kNN	0.678	0.660	0.889	0.662	0.110	0.660	0.551
MLP	0.666	0.667	0.885	0.643	0.114	0.637	0.525
(b) mFr_stroke dataset ($\hat{X}$)							
Algorithms	Accuracy rate	Sensitivity	Specificity	Precision	False Positive Rate	F1- score	MCC
Decision Tree	0.860	0.850	0.952	0.846	0.047	0.848	0.801
Naïve Bayes	0.864	0.877	0.955	0.844	0.044	0.855	0.813
kNN	0.513	0.492	0.832	0.466	0.167	0.474	0.310
MLP	**0.886**	0.837	0.961	0.868	0.038	0.871	0.834



(b) Confusion matrix results for four stroke subtypes in mFr_stroke dataset $(\hat{X})$ with (i) Decision Tree (ii) Naive Bayes (iii) kNN (iv) MLP

Fig. 3. (Figure 2. Cont.) Confusion matrix results for four stroke subtypes in stroke dataset (X) and mFr_stroke dataset $(\hat{X})$ using (i) Decision tree (ii) Naïve Bayes (iii) k-NN (iv) MLP.

Step 3: The accuracy rate results based on decision tree, Naïve bayes, kNN and MLP algorithms for the classification of stroke subtypes on stroke dataset (X) and mFr_stroke dataset $(\hat{X})$.

In this study, Decision tree, Naïve bayes, kNN and MLP algorithms were applied to the stroke dataset (X) and mFr_stroke dataset $(\hat{X})$ for the classification of stroke subtypes. Accordingly, Table 3 shows the accuracy rates (Error, Sensitivity, Specificity, Precision, False Positive Rate, F1-score and Matthews Correlation Coefficient (MCC)) of the aforementioned algorithms for four subtypes of stroke.

As Table 3 shows, the highest accuracy rate (0.886) has been obtained from the mFr_stroke dataset $(\hat{X})$ to which multifractal regularization method (Kullback norm) with the MLP algorithm. In addition, for Naïve Bayes and K-nn, the mFr_stroke dataset $(\hat{X})$ also yielded better results compared to those of stroke dataset (X). The experimental results of our study demonstrate the critical and determining role of multifractal regularization method (Kullback norm) in terms of accuracy based on classification performance.

4 Conclusion

The principal contribution of the study is related to the self-similarity and regularity detection in stroke dataset (X). Since brain displays highly complex, regular as well as self-similar patterns, its analysis requires both morphological and hierarchical views. Based on this view, the method we proposed in this study manifests novelties, through two different approaches, as obtained from the experimental results when compared with previous works [17,30]. As the first approach, the application of decision tree, Naïve bayes, kNN and MLP algorithms was applied to the stroke dataset, which is one of the novelties in terms of the classification of extensive stroke subtypes dataset. b) The second approach incorporates two stages, the first one being multifractal regularization (kulback normalization) as applied to the stroke dataset (X), and mFr_stroke dataset ($\hat{X}$) having been generated subsequently. The second stage is the application of the four algorithms to the mFr_stroke dataset ($\hat{X}$) as per accuracy, specificity, sensitivity, precision, F1-score and Matthews Correlation Coefficient (MCC). The application of these methods in this way can be the second novel contribution of this study. It is the first time multifractal regularization method has been applied to the stroke dataset's subtypes. The experimental results of our study demonstrate the critical and determining role of multifractal regularization method (Kullback norm) in terms of accuracy based on classification performance since the highest accuracy rate was obtained from mFr_stroke dataset ($\hat{X}$). All things considered, the current study has demonstrated the reliability and applicability with the methods proposed and conducted. Thus, the integrated method utilized has revealed the significance of multifractal patterns and accurate prediction of diseases in diagnostic and other critical-decision making processes in related fields. Accordingly, the integrated approach of this study can serve a facilitating purpose to the physicians regarding the diagnostic and predictive processes for diseases such as stroke and other medical incidents that have to do with the complex dynamics and structures like that of brain.

References

1. Roca, J.L., Rodrıguez-Bermudez, G., Fernandez-Martinez, M.: Fractal-based techniques for physiological time series: an updated approach. Open Phys. **16**(1), 741–750 (2018)
2. Di Ieva, A.: The Fractal Geometry of the Brain, vol. 585. Springer, New York (2016)
3. Li, Z., Liu, Z., Khan, M.A.: Fractional investigation of bank data with fractal-fractional Caputo derivative. Chaos, Solitons Fractals **131**, 109528 (2020)
4. Karaca, Y., Zhang, Y.D., Muhammad, K.: Characterizing complexity and self-similarity based on fractal and entropy analyses for stock market forecast modelling. Expert Syst. Appl. **144**, 113098 (2020)
5. Karaca, Y., Cattani, C.: A comparison of two hölder regularity functions to forecast stock indices by ANN algorithms. In: Misra, S., et al. (eds.) ICCSA 2019. LNCS, vol. 11620, pp. 270–284. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-24296-1_23

6. Dimri, V.P., Ganguli, S.S.: Fractal theory and its implication for acquisition, processing and interpretation (API) of geophysical investigation: a review. J. Geol. Soc. India **93**(2), 142–152 (2019)
7. Nottale, L.: Scale relativity and fractal space-time: theory and applications. Found. Sci. **15**(2), 101–152 (2010)
8. Petrica, V., Maricel, A.: On the transport phenomena in composite materials using the fractal space-time theory. Adv. Compos. Mater. Med. Nanotechnol. **477**, 477–494 (2011)
9. Meltzer, M.I.: The potential use of fractals in epidemiology. Prev. Vet. Med. **11**(3–4), 255–260 (1991)
10. Levy-Vehel, J.: Fractal approaches in signal processing. Fractals **3**(04), 755–775 (1995)
11. Albertovich, T.D., Aleksandrovna, R.I.: The fractal analysis of the images and signals in medical diagnostics. Fract. Anal. Appl. Health Sci. Soc. Sci. **26**, 57 (2017)
12. Karaca, Y., Cattani, C.: Clustering multiple sclerosis subgroups with multifractal methods and self-organizing map algorithm. Fractals **25**(04), 1740001 (2017)
13. Karaca, Y., Moonis, M., Baleanu, D.: Fractal and multifractional-based predictive optimization model for stroke subtypes' classification. Chaos, Solitons Fractals **136**, 109820 (2020)
14. Karaca, Y., Cattani, C., Moonis, M., Bayrak, Ş.: Stroke subtype clustering by multifractal Bayesian denoising with fuzzy C means and K-means algorithms. Complexity **2018**, 15 pages (2018). Article ID 9034647
15. Karaca, Y., Moonis, M., Zhang, Y.D., Gezgez, C.: Mobile cloud computing based stroke healthcare system. Int. J. Inf. Manag. **45**, 250–261 (2019)
16. Norrving, B.: Oxford Textbook of Stroke and Cerebrovascular Disease. Oxford University Press, Oxford (2014)
17. He, L., Wang, J., Zhang, L., Zhang, X., Dong, W., Yang, H.: Decreased fractal dimension of heart rate variability is associated with early neurological deterioration and recurrent ischemic stroke after acute ischemic stroke. J. Neurol. Sci. **396**, 42–47 (2019)
18. Smitha, B.: Fractal and multifractal analysis of atherosclerotic plaque in ultrasound images of the carotid artery. Chaos, Solitons Fractals **123**, 91–100 (2019)
19. Lemmens, S., Devulder, A., Van Keer, K., Bierkens, J., De Boever, P., Stalmans, I.: Systematic review on fractal dimension of the retinal vasculature in neurodegeneration and stroke: assessment of a potential biomarker. Front. Neurosci. **14**, 16 (2020)
20. Appaji, A., et al.: Retinal vascular fractal dimension in bipolar disorder and schizophrenia. J. Affect. Disord. **259**, 98–103 (2019)
21. Amezquita-Sanchez, J.P., Mammone, N., Morabito, F.C., Marino, S., Adeli, H.: A novel methodology for automated differential diagnosis of mild cognitive impairment and the Alzheimer's disease using EEG signals. J. Neurosci. Methods **322**, 88–95 (2019)
22. Zhao, P., Zhuo, R., Li, S., Lin, H., Shu, C.M., Laiwang, B., Suo, L.: Fractal characteristics of gas migration channels at different mining heights. Fuel **271**, 117479 (2020)
23. Zuo, X., Tang, X., Zhou, Y.: Influence of sampling length on estimated fractal dimension of surface profile. Chaos, Solitons Fractals **135**, 109755 (2020)
24. Razminia, K., Razminia, A., Shiryaev, V.I.: Application of fractal geometry to describe reservoirs with complex structures. Commun. Nonlinear Sci. Numer. Simul. **82**, 105068 (2020)

25. Andres, J., Langer, J., Matlach, V.: Fractal-based analysis of sign language. Commun. Nonlinear Sci. Numer. Simul. **84**, 105214 (2020)
26. Raghavendra, U., Acharya, U.R., Adeli, H.: Artificial intelligence techniques for automated diagnosis of neurological disorders. Eur. Neurol. **82**, 41–64 (2019)
27. Cheon, S., Kim, J., Lim, J.: The use of deep learning to predict stroke patient mortality. Int. J. Environ. Res. Public Health **16**(11), 1876 (2019)
28. Ge, Y., et al.: Predicting post-stroke pneumonia using deep neural network approaches. Int. J. Med. Inf. **132**, 103986 (2019)
29. Lin, C.H., et al.: Evaluation of machine learning methods to stroke outcome prediction using a nationwide disease registry. Comput. Methods Programs Biomed. **190**, 105381 (2020)
30. Arai, K., Bhatia, R., Kapoor, S. (eds.): CompCom 2019. AISC, vol. 997. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-22871-2
31. The MathWorks. MATLAB (R2019b) The mathWorks, inc., Natick, MA (2019)
32. Vehel, L., FracLab (2019) project.inria.fr/fraclab/
33. Harte, D.: Multifractals. Chapman and Hall, London (2001)
34. Lopes, R., Betrouni, N.: Fractal and multifractal analysis: a review. Med. Image Anal. **13**(4), 634–649 (2009)
35. Moreno, P.A., et al.: The human genome: a multifractal analysis. BMC Genom. **12**, 506 (2011)
36. Barnsley, M.F.S., Saupe, D., Vrscay, E.R.: Signal enhancement based on hölder regularity analysis. IMA Vol. Math. Appl. **132**, 197–209 (2002)
37. Ben Slimane, M., Ben Omrane, I., Ben Abid, M., Halouani, B., Alshormani, F.: Directional multifractal analysis in the L^p setting. J. Funct. Spaces **2019**, 12 pages (2019). Article ID 1691903
38. Levy Vehel, J.: Signal enhancement based on Hölder regularity analysis. Inria technical report (1999)
39. Heurteaux, Y., Jaffard, S.: Multifractal analysis of images: new connexions between analysis and geometry. In: Byrnes, J. (ed.) Imaging for Detection and Identification, pp. 169–194. Springer, Dordrecht (2007). https://doi.org/10.1007/978-1-4020-5620-8_9
40. Jaffard, S.: Pointwise regularity criteria. C.R. Math. **339**(11), 757–762 (2004)
41. Jaffard, S. and Melot, C.: Wavelet analysis of fractal boundaries. Part 1: local exponents. Commun. Math. Phys. **258**(3), 513–539 (2005)
42. Ben Slimane, M. and Mélot, C.: Analysis of a fractal boundary: the graph of the knopp function. Abstract Appl. Anal. **2015** 14 (2015). Article number 587347
43. Shao, J., Buldyrev, S.V., Cohen, R., Kitsak, M., Havlin, S., Stanley, H.E.: Fractal boundaries of complex networks. EPL (Europhys. Lett.) **84**(4), 48004 (2008)
44. Lutton, E., Grenier, P., Vehel, J.L.: An interactive EA for multifractal Bayesian denoising. In: Rothlauf, F., et al. (eds.) EvoWorkshops 2005. LNCS, vol. 3449, pp. 274–283. Springer, Heidelberg (2005). https://doi.org/10.1007/978-3-540-32003-6_28
45. Donoho, D.L.: De-noising by soft-thresholding. IEEE Trans. Inf. Theory **41**(3), 613–627 (1994)
46. Mandelbrot, B.B., Van Ness, J.W.: Fractional Brownian motions, fractional noises and applications. SIAM Rev. **10**(4), 422–437 (1968)
47. Chen, Y.P., Chen, Y., Tong, L.: Sar Image Denoising Based on Multifractal and Regularity Analysis. In Key Engineering Materials. Trans Tech Publications Ltd. 500, 534–539 (2012)

48. Karaca, Y., Cattani, C.: Computational methods for data analysis. Walter de Gruyter GmbH, Berlin (2018)
49. Zhang, Y., Sun, Y., Phillips, P., Liu, G., Zhou, X., Wang, S.: A multilayer perceptron based smart pathological brain detection system by fractional Fourier entropy. J. Med. Syst. **40**(7), 173 (2016)