WILEY | Hindawi

Research Article

Stroke Subtype Clustering by Multifractal Bayesian Denoising with Fuzzy C Means and K-Means Algorithms

Yeliz Karaca [iD],[1] Carlo Cattani,[2] Majaz Moonis,[3] and Şengül Bayrak [iD][4]

[1] *University of Massachusetts Medical School, Worcester, MA 01655, USA*
[2] *Engineering School, DEIM, University of Tuscia, 01100 Viterbo, Italy*
[3] *Department of Neurology and Psychiatry, University of Massachusetts Medical School, Worcester, MA 01655, USA*
[4] *Department of Computer Engineering, Haliç University, 34000 İstanbul, Turkey*

Correspondence should be addressed to Yeliz Karaca; yeliz.karaca@ieee.org

Multifractal denoising techniques capture interest in biomedicine, economy, and signal and image processing. Regarding stroke data there are subtle details not easily detectable by eye physicians. For the stroke subtypes diagnosis, details are important due to including hidden information concerning the possible existence of medical history, laboratory results, and treatment details. Recently, K-means and fuzzy C means (FCM) algorithms have been applied in literature with many datasets. We present efficient clustering algorithms to eliminate irregularities for a given set of stroke dataset using 2D multifractal denoising techniques (Bayesian (mBd), Nonlinear (mNold), and Pumping (mPumpD)). Contrary to previous methods, our method embraces the following assets: (a) not applying the reduction of the stroke datasets' attributes, leading to an efficient clustering comparison of stroke subtypes with the resulting attributes; (b) detecting attributes that eliminate "insignificant" irregularities while keeping "meaningful" singularities; (c) yielding successful clustering accuracy performance for enhancing stroke data qualities. Therefore, our study is a comprehensive comparative study with stroke datasets obtained from 2D multifractal denoised techniques applied for K-means and FCM clustering algorithms. Having been done for the first time in literature, 2D mBd technique, as revealed by results, is the most successful feature descriptor in each stroke subtype dataset regarding the mentioned algorithms' accuracy rates.

1. Introduction

Multifractal analysis is concerned with the study of the regularity structure of processes, both from a local and global point of view. Multifractal Bayesian denoising is a technique on regularity-based enhancement and it acts by finding data that is close to the observations along with the multifractal data prescribed. This method depends on the tuning of a small set of parameters which are capable of providing different improvements pertaining to the observed noisy data. In many applications, this has been successfully utilized in cases in which irregularity carries important information. Abundant natural phenomena in fields like physics, finance, construction, environment, medicine, and biology have been shown to display a fractal behavior [1–3].

Being the third most frequent cause of death following heart disease and cancer in developed countries, stroke is among the most common causes of cognitive impairment and vascular dementia [4]. Stroke can be described as the quick loss of brain function owing to the disturbance occurring in the blood supply to the brain [5]. Being one of the foremost causes of death worldwide, stroke can be generally classified into two types. The clinical course types are defined as the ischemic stroke and hemorrhagic stroke [6].

In this study, we worked on the dataset of individuals who have been diagnoses with ischemic stroke: no stroke/TIA, large vessel, small vessel, cardioembolic, cryptogenic, dissection, other (moyamoya, FMD, hereditary, coagulopathy, vasculitis, other rare). No stroke/TIA is transient ischemic attack; it is a mini stroke or a minor stroke that happens

when a blood clot blocks an artery for a brief period of time [7]. Large vessel is infarction happens owing to artery-to-artery or low flow embolism in the ipsilateral arterial tree (intracranial or extracranial) segments of carotid or vertebrobasilar arteries, or proximal middle cerebral artery [8]. Small vessel diseases of the cerebral vasculature contribute to varied forms of brain dysfunction cell death and injury. Small vessel disease of the brain corresponds to ≈25% to 30% of strokes, and it is a primary cause of cognitive decline and disability due to age-related and hypertension-related reasons [9]. Cardioembolic stroke is mainly preventable, calling for efforts at primary prevention for major-risk cardioembolic sources. When stroke as a result of cardiac embolism has happened, the chances of recurrence are comparatively high for most cardioembolic sources. In addition, cryptogenic stroke is commonly seen in clinical practice. Cryptogenic stroke is defined as a brain infarction not attributable to a source of definite cardioembolism, small artery disease, or large artery atherosclerosis, even with a standard cardiac, vascular, and serologic evaluation [10]. Dissection stroke is cervical arterial dissection; it is a main reason for stroke experienced in young adults. However, the management thereof remains uncertain despite the standard treatment administered through anticoagulants or antiplatelet drugs [11]. Demographic information, medical history, results of the laboratory tests, treatments, and medications are among the most important data belonging to the patients, which can be categorized into the following major groups.

In recent years, the fractal and multifractal analysis in biomedical data has seen a growing interest. Regarding the relevant topics, in the following studies, Wang et al. [12], Yang et al. [13], Karaca and Cattani [14], Tsaneva [15], Doubal et al. [16], Shanmugavadivu et al. [17], and Ahammer et al. [18] underlined the significance of fractal and multifractal techniques for data analysis in medicine. It has also been acknowledged that the multifractal techniques have successful feature descriptor in stroke applications [19–21]. Numerous studies reveal successful clustering results regarding the stroke subtypes with applications of K-means and FCM [22, 23]. However, it has been also seen that there is a shortage in literature and subject matter as to studies with combined applications of numeric data [24, 25], 2D multifractal denoising techniques, and machine learning approaches.

Our approach aims to be broader and more complete since this study of ours is large and comprehensive when compared with other studies done with stroke dataset [22, 23] in literature, taking into consideration the dimension of 2204 (the number of patients with 7 different stroke subtypes) and 23 attributes. The 7 different stroke subtypes are as follows: no stroke/TIA, large vessel, small vessel, cardioembolic, cryptogenic, dissection, other (moyamoya, FMD, hereditary, coagulopathy, vasculitis, other rare). The attributes include demographic information, medical history, results of laboratory tests, treatments, and medications. The clustering of the subtypes of stroke is a remarkable challenge in its own term. Besides, all researches have been done on many different kinds of analysis regarding stroke dataset, but no work has been reported yet which relates attributes (demographic information, medical history, results of laboratory tests,
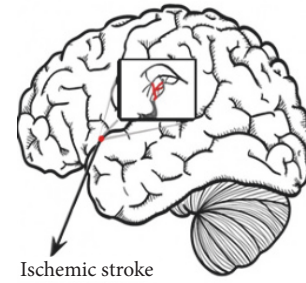

Ischemic stroke

FIGURE 1: Ischemic stroke image.

treatments, and medications) through the 2D multifractal denoising techniques to fuzzy C means and K-means algorithms applied for clustering purposes. For this reason, 2D multifractal denoising techniques (mBd, mNold, mPumpD) have been administered for the identification of significant and efficient attributes belonging to the patients (among 23 of them) for the clustering of 7 subtypes of stroke. It is adapted well to the case in which the data to be recovered is very irregular and nowhere differentiable, a property relevant to fractal or self-similar structures. We obtained regularity-based enhancement from 2D multifractal denoising techniques datasets (mBd, mNold, mPumpD). These datasets are clustered using the K-means and FCM algorithms. 2D mBd stroke dataset has yielded better clustering than stroke dataset, 2D mNold stroke dataset, and 2D mPumpD stroke dataset of stroke subtypes. When compared with studies mentioned above, our study is a comprehensive and comparative one since the stroke datasets as obtained from 2D multifractal denoised techniques have been applied for the first time in literature for K-means and FCM clustering algorithms.

The paper is organized as follows: Section 2 provides Materials and Methods. Methods of our approach are basic facts on Hölder regularity and multifractal analysis, multifractal Bayesian denoising in $\mathcal{S}(g, \psi)$, numerical experiments (stroke dataset experiments in multifractal Bayesian denoising technique), K-means algorithm, and fuzzy C means algorithm. As the last sections, results and discussion and conclusions are presented in Sections 3 and 4, respectively.

2. Materials and Methods

2.1. Patient Details. 2204 individuals from Massachusetts Medical School, University of Worcester, Massachusetts, USA, were kept under observation in this study. Data were collected in the period between March 9, 2007, and October 2, 2016. The individuals had ischemic stroke diagnosis (see Figure 1). The ischemic strokes in the dominant hemisphere lead to more functional deficits compared to the strokes in the nondominant hemisphere as they are evaluated on the National Institutes of Health Stroke Scale (NIHSS).

A total of 2204 patients (414 males [labelled with (1)] and 1790 females [labelled with (0)]) were included in our experiments. Stroke patients are aged between 0 and 104, with seven subtypes of ischemic stroke being examined in this study.

TABLE 1: Breakdown of stroke patients by age.

Stroke subtypes	Age					
	0–29	30–36	37–49	50–70	71–90	91–104
No stroke/TIA	7	2	13	62	72	11
Large vessel	18	2	17	216	208	20
Small vessel	14	1	7	113	85	8
Cardioembolic	18	5	21	195	391	59
Cryptogenic	22	12	53	228	196	17
Dissection	2	2	14	30	8	3
Other	5	5	7	17	17	1

In this study, demographic information, medical history, results of laboratory tests, treatments, and medications data, as can be seen in Table 1, pertaining to 2204 stroke subtypes patients. Table 2 provides the main headings of attributes used for the stroke subtypes.

Baseline characteristics of the patients involved as stratified by infarct side are outlined in Table 3 regarding the stroke dataset.

Medications which are given to the patients are classified into broad categories: statin, antiHTN, antidiabetic, antiplatelet, anticoagulation. Attributes of CT perfusion and neurointervention are utilized for the treatment. The modified Rankin Scale (mRS) score was evaluated at 90 days by a physician who had training on strokes or a stroke nurse with knowledge of strokes and certified in mRS via in-person or via phone interview. We comply with the Strengthening the Reporting of Observational Studies in Epidemiology guideline (https://www.strobe-statement.org/). 0.09% of the individuals' disorder progressed to the hemorrhagic stroke.

2.2. Methods. In this study, we have provided two potential contributions. We introduced the 2D mBd, 2D mNold, and 2D mPumpD which are relatively novel multifractal techniques calculating the regular data from stroke data. We proposed the use of stroke dataset and regular and denoised stroke datasets (2D mBd, 2D mNold, 2D mPumpD) to be trained with unsupervised K-means and FCM algorithms for the clustering along with the aim of improving the clustering performance of the stroke subtypes. Our method is reliant on the steps specified below:

(a) 2D multifractal denoising techniques (2D mBd, 2D mNold, 2D mPumpD) were applied to the stroke dataset (which can be seen in Table 2). In order to identify the stroke dataset significant regularity, which is the fundamental concept concerning multifractal denoising; the best is explained on a simple example. The aim is to eliminate "insignificant" irregularities while retaining "meaningful" singularities and denoised dataset.

(b) The stroke datasets obtained from stroke dataset and 2D multifractal denoising techniques (mBd, mNold, and mPumpD) were clustered by having been applied to the K-means and FCM algorithms.

(c) The comparisons of datasets (stroke dataset, mBd stroke dataset, mNold stroke dataset, and mPumpD stroke dataset) were performed with the K-means and FCM algorithms as to the clustering accuracies. Computations and figures were obtained by Matlab, Mosek, and FracLab [26] environment.

2.2.1. Basic Facts on Hölder Regularity and Multifractal Analysis. We are concerned with the enhancement, or denoising, of complex data, which is the stroke dataset, relying on the analysis of the local Hölder regularity. We rather suppose that data enhancement is comparable to increasing the Hölder regularity at every point [27]. Such methods are adapted well to the case in which the data which would be recovered is highly irregular, for instance, nowhere differentiable with local regularity that varies rapidly.

In this paper, our focus is on the pointwise Hölder exponent for simplifying the notations, and we assume that our data are not differentiable [28, 29].

Let $\alpha \in (0, 1)$ and $x_0 \in Q \subset \mathfrak{R}$. A function $f : Q \to \mathfrak{R}$ is in $L^{\alpha}_{x_0}$ if, for all x in a neighborhood of x_0,

$$\left| f(x) - f(x_0) \right| \le l \left| x - x_0 \right|^{\alpha}, \tag{1}$$

where l is a constant. Pointwise Hölder exponent of f at x_0, denoted by $\alpha(x_0)$, is the supremum of the α for which (1) is valid.

In this paper, we shall concentrate on the statistical approach that brings about consideration of a quantity named the *large deviation multifractal spectrum* [29, 30]. This spectrum can be defined as follows:

Consider a stochastic process $X(t)$, $t \in T \subset R$ on a probability space (Ω, F, P). For convenience in terms of notation, we will assume without loss of generality that $T = [0, 1]$.

Set $N^{\varepsilon}_n(\alpha) = \#\{q : \alpha - \varepsilon \le \alpha^q_n \le \alpha + \varepsilon\}$, where α^q_n is the coarse-grained Hölder exponent that corresponds to the dyadic interval $I^q_n = [q2^{-n}, (q + 1)2^{-n}]$; that is, $\alpha^q_n = \log |Y^q_n| / - \log n$. At this point, (#)is the number of nonempty boxes, and its measure is portrayed by exponents from the interval $(\alpha - \varepsilon, \alpha + \varepsilon)$ [29]. At this point, Y^q_n is some quantity measuring the variation of X in the interval I^q_n. The choice $Y^q_n := X((q + 1)2^{-n}) - X(q2^{-n})$ brings about the simplest analytical computations. Another possibility, which shall be the one used in this paper, is to take Y^q_n to be the $x_{n,q}$ of X at scale n and location q [29]. This definition is convenient in various aspects, since it ensures the utilization of the wavelet bases' versatility. However, it also has a setback: As a matter of fact, the multifractal spectrum obtained accordingly will rely predominantly on the wavelet chosen ψ. Hence, if one sets $Y^q_n : x_{n,q}$, it would not make sense to speak of the spectrum of X without reference to the analyzing wavelet chosen. Along the paper, the wavelet coefficient of data X is denoted by $x_{j,q}$, with j being the scale and q being the location [29, 30].

Then coarse-grained multifractal large deviation spectrum is provided by the equation $f_g(\alpha) = \lim_{\varepsilon \to 0} \lim_{n \to \infty} \sup(\log N^{\varepsilon}_n / \log n)$. The definition of $f_g(\alpha)$ is connected with the large deviation theorem that provided the probabilistic interpretation for the multifractal spectrum.

TABLE 2: Stroke dataset.

Number of stroke subtypes/TOAST	Main heading of attributes	Data size
No stroke/TIA (167) Large vessel (481) Small vessel (228) Cardioembolic (689) Cryptogenic (528) Dissection (59) Other (52)	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, neurointervention)	2204×23

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 artery 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.

TABLE 3: Stroke dataset description.

Attributes	Status	Number of patients/values (%)	Descriptions
HTN	Yes	1593 (72%)	Hypertension
Hyperlip	Yes	1197 (54%)	High levels of lipid (fat) in blood
DM	Yes	602 (27%)	Diabetes
H/O stroke/TIA	Yes	546 (25%)	History of stroke/TIA
AttrialFib	Yes	541 (25%)	Abnormal heart rhythm
CAD	Yes	513 (23%)	Coronary artery disease
CHF	Yes	229 (10%)	Congestive heart failure
PAD/carotid Disease	Yes	318 (14%)	Peripheral artery disease
Tobacco	Yes	520 (23%)	Cigarette addict
ETOH	Yes	308 (1.7%)	Alcohol addict
Statin	Yes	1000 (45%)	Medications given to the patient are grouped into five broad categories
AntiHTN	Yes	1332 (60%)	
Antidiabetic	Yes	454 (20%)	
Antiplatelet	Yes	1031 (47%)	
Anticoagulation	Yes	242 (10%)	
CT perfusion	Yes	137 (0.06%)	Procedures used for treatment
Neurointervention	Yes	271 (12%)	
mRS 90 days	Low	2007	Dichotomized into low (0–2), high (3–6)
	High	197	
Hemorrhagic con.	Yes	204 (0.09%)	Whether the ischemic stroke turned to hemorrhagic
NIHSS admission		9.3 +/− 8.3	Measures the severity of stroke
TPA	Yes	413 (19%)	TPA (tissue plasminogen activator) is used to break down blood clots

It should be noted that irrespective of the choice of Y_n^q, f_g all the time ranges in $\Re^+ \cup \{-\infty\}$. The value $-\infty$ corresponds to values of the coarse-grained exponent which are not observed at all sufficiently on small scaler [29].

The intuitive meaning of f_g can be found as follows: For n large enough, one has approximately $P_n(\alpha_n^q \simeq \alpha) \simeq 2^{-n(1-f_g(\alpha))}$, in which P_n denote the uniform distribution over $\{0, 1, \ldots, 2^n - 1\}$. Hence, for all α such that $f_g < 1$, $1 - f_g(\alpha)$

measures the exponential rate of decay of the probability of finding an interval I_n^q with coarse-grained exponent equal to α, when n tends to infinity. As a whole, f_g is a random function. In the applications, it is convenient to regard the following deterministic version of f_g:

$$F_g(\alpha) = 1 + \lim_{\varepsilon \to 0} \lim_{n \to \infty} \sup \frac{\log \pi_n^\varepsilon(\alpha)}{\log(n)} \tag{2}$$

$$\text{where } \pi_n^\varepsilon(\alpha) := P \times P_n\left[\alpha_n^q \in (\alpha - \varepsilon, \alpha + \varepsilon)\right].$$

See [29].

Here, we present technique based on the multifractal data instead of applying it to the use of the Hölder exponent merely. This approach can generally ensure more robust estimates because a higher level description is used for subsuming information on the entire data. Besides this, we also assume a semiparametric approach. To put it more specifically, it can be said that we put forth the assumption that the considered data belong to a given set of parameterized classes and are explained currently [31].

Let us state that F is the set of lower semicontinuous functions from $\mathfrak{R}^+$ to $\mathfrak{R}^+ \cup \{-\infty\}$ [29]. We regard classes of random functions as $X(t)$, $t \in [0, 1]$, which can be defined as (Ω, F, P) and described according to Definition 1 which can be seen below. Each cluster $\mathcal{S}(g, \psi)$ is considered by the functional parameter $g \in F$ and a wavelet ψ. And the set $\{\psi_{j,q}\}_{j,q}$ constitutes a basis of L^2. Let Q be a positive constant, as defined based on

$$P_j^\varepsilon(\alpha, Q) = P \times P_j\left(\alpha - \varepsilon < \frac{\log_2(Q|x_{j,q}|)}{-j} < \alpha + \varepsilon\right). \tag{3}$$

Definition 1. $\mathcal{S}(g, \psi) = \{X : \exists Q > 0, j_0 \in Z : \forall j > j_0, x_{j,q}, x_{j,q'} \text{ are distributed identically for } (q, q') \in \{0, 1, \ldots, 2^j - 1\} \text{ and } \log_2 P_j^\varepsilon(\alpha, Q)/j = g(\alpha) + \mathfrak{R}_{n,\varepsilon}(\alpha)\}$ in which $\mathfrak{R}_{n,\varepsilon}(\alpha)$ is $\lim_{\varepsilon \to 0} \lim_{n \to \infty} \mathfrak{R}_{n,\varepsilon}(\alpha) = 0$ in a uniform way in α.

For j large sufficiently, the assumption that the wavelet coefficients $(x_{j,q})_q$ at scale j are distributed identically requires the following:

$$\pi_n^\varepsilon(\alpha) := P \times P_n\left[\alpha_n^q \in (a - \varepsilon, a + \varepsilon)\right],$$

$$\pi_j^\varepsilon(\alpha, Q) := P \times P_j\left(\alpha - \varepsilon < \frac{\log_2(Q|x_{j,q}|)}{-j}\right) < \alpha + \varepsilon \tag{4}$$

$$= P\left(\alpha - \varepsilon < \frac{\log_2(Q|x_{j,q}|)}{-j} < \alpha + \varepsilon\right).$$

Consequently, Definition 1 yields a plain interpretation regarding multifractal analysis. We consider the set of random data X for a given wavelet ψ. Accordingly, the normalized data QX has deterministic multifractal data $F_g(\alpha)$ that equals $1 + g$ along with the following further condition: F_g is gained as a limit in j lim sup and this limit has been attained in a uniform manner into α. This condition confirms that the rescaled statistics of the $\alpha_{j,q}$ are sufficiently close to their limit for large enough j, which allows a meaningful inference. The cluster $\mathcal{S}(g, \psi)$ includes a wide variety of data.

2.2.2. Multifractal Bayesian Denoising in $\mathcal{S}(g, \psi)$. At this point, the key steps in the classical Maximum A Posteriori (MAP) approach in a Bayesian frame can be recalled, as adjusted to our setting. It is observed that the noisy data is Y, and it is assumed that $Y = X + B$, where B is a noise independent from original data X, with known law as B (it should be noted that we use orthonormal wavelets; B is denoted in equation as b). Hence, we have $y_{j,q} = x_{j,q} + b_{j,q}$. The map estimate $\widehat{x}_{j,q}$ of $x_{j,q}$ from the observation $y_{j,q}$ is defined to be an argument maximizing $P(x_{j,q}/y_{j,q})$. Since $P(y_{j,q})$ does not depend on $x_{j,q}$, using Bayes rules and maximizing $P(x_{j,q}/y_{j,q})$ correspond to maximizing the product $P(y_{j,q}/x_{j,q})P(x_{j,q})$ [29, 30].

The MAP estimate [30] is stated as $\widehat{x}_{j,q} = \arg\max_x[P(y_{j,q}/x)P(x)]$. The term $P(y_{j,q}/x)$ can be calculated from the law of B easily if B is assumed to be (0), since $b_{j,q}$ share the same law as B. It can also be recalled that orthonormal wavelets are used by us. The preceding $P(x_{j,q})$ is inferred from our assumption that X belongs to $\mathcal{S}(g, \psi)$ as follows: for $x > 0$, set $\alpha_j(x) = \log_2(Qx)/-j$,

$$P\left(\left|x_{j,q}\right| = x\right) = P\left(\frac{\log_2(Q|x_{j,q}|)}{-j} = \alpha_j(x)\right) \tag{5}$$

$$\simeq 2^{j(g(\alpha_j(x))-1)}.$$

This leads to the definition of an approximate Bayesian MAP estimate as

$$\widehat{x}_{j,q} = \arg\max_{x>0}\left[jg\left(\frac{\log_2(\widehat{Q}x)}{-j} + \log_2\left(P\left(\frac{y_{j,q}}{x}\right)\right)\right)\right] \tag{6}$$

$$\cdot \text{sgn}(y_{j,q}),$$

where $\text{sgn}(y)$ is the sign of y and $\widehat{Q} = (\sup_{j>j_0} \sup_q(x_{j,q}))^{-1}$. The estimate for Q can be justified in a heuristic way as follows: $\log_2(Q|x_{j,q}|)/-j \simeq \alpha$ with $\alpha > 0$ suggesting that $Q|x_{j,q}| < 1$ for all the couples (j, q). $\widehat{Q}$ is chosen as the smallest normalizing factor requiring the latter inequality. In our experiments, we address the incident in which the noise is centered, Gaussian, with variance σ^2. The MAP estimate can be seen in accordance with

$$\widehat{x}_{j,q} = \arg\max_{x>0}\left[jg\left(\frac{\log_2(\widehat{Q}x)}{-j} - \frac{(y_{j,q} - x)^2}{2\sigma^2}\right)\right] \tag{7}$$

$$\cdot \text{sgn}(y_{j,q}).$$

Equation (6) provides an explicit formula for Y denoising, but it often offers limited practical use. In fact, one is not aware of the multifractal data of X in most of the applications. If there is no evaluation of g, *it* would not be probable to use (6) for the purpose of obtaining $\widehat{x}_{j,q}$. Moreover, one should bear in mind that in general F_g depends on the analyzing wavelet. Hence, it is necessary to understand the shape of the data for a given wavelet. Moreover, the main goal of our approach is to remove the multifractal characteristics

of X from the denoised data $\widehat{X}$: regarding the Multifractal Bayesian approach use in our study, a strong justification is to be able to estimate F_g^X in the following manner: (a) denoise Y, (b) evaluate the data $F_g^{\widehat{X}}$ numerically, (c) set $\widehat{F}_g^X = F_g^{\widehat{X}}$. It will be obvious from this approach that it does not seem right to have the necessity of previous knowledge of F_g^X in the Bayesian approach. Hence, we present a "degenerated" version of (6), and here the input is used as a single real parameter rather than the whole data. The heuristic reads as follows: from regularity perspective a significant piece of information in the data is concerned with its support, for example, the set of all the occurring Hölder exponents. α_0 denotes the smallest regularity observed actually in the data [32, 33]. The shapes of the F_g spectra gained through different analyzing wavelets is reliant on the wavelet but their support is included in $[\alpha_0, \infty)$. The "flat data," therefore, encompasses inherent information. It only relies on the positive real α_0. Rewriting (6) with a flat data gives the explicit simple expression as in

$$\widehat{x}_{j,q} = \begin{cases} y_{j,q} & \text{if } Q\left|y_{j,q}\right| < 2^{-j\alpha_0} \\ 2^{-j\alpha_0} & \text{otherwise.} \end{cases} \tag{8}$$

α_0 is really a priori information but it is possible to be predicted from the noisy observations. In view of that, it can be analogous to the threshold that is used in the classical soft or hard wavelet thresholding scheme. It would be beneficial to regard α_0 as a tuning parameter in the applications. Increasing α_0 offers a more smooth estimate (as it is assumed that the original data have a larger minimal exponent). It would be remarkable to make a comparison with the hard-thresholding policy on the wavelet coefficients (see more details in [29]).

2.2.3. Numerical Experiments.

We present some results regarding the stroke dataset. In each case, the result of the Bayesian multifractal denoising and the classical hard-thresholding technique is shown. For all procedures and stroke dataset, the parameters (see Tables 2 and 3) are set in order to obtain the best fit to the known original data. On the whole, the following conclusions can be inferred from these experiments. It is observed that for the irregular data, like the ones handled here, which belong to $\mathcal{S}(g, \psi)$, the Bayesian method yields more satisfactory results compared to those of classical wavelet thresholding. This method particularly preserves a roughly correct regularity along the path, whereas the wavelet shrinkage yields data that have both too smooth and too irregular regions. In this study, numerical experiments were obtained with 2D mBd, 2D mNold, and 2D mPumpD techniques being applied to numerical experiments pertaining to the stroke dataset (see Table 2).

(a) Stroke Dataset Experiments in Multifractal Bayesian Denoising Technique. Stroke dataset (X) (see Table 2) was applied on FracLab [26] program to regularity-based enhancement $(\widehat{X})$ from 2D multifractal denoising techniques.

The steps of multifractal Bayesian denoising technique applied on the stroke dataset to be able to get the regular and denoised stroke dataset are provided below.

Step 1. We consider

$$\mathcal{S} = \left\{ X : \left|x_{j,q}\right| = Q\epsilon_{j,q} 2^{-j\gamma} \right\}. \tag{9}$$

Step 2. For each attribute in the stroke dataset $(X = x_{j=1,\dots,2204 \times 23})$ MAP $(\widehat{x}_{j,q}$ of $x_{j,q})$ values are calculated. Here, Q and γ are the positive constants and $\varepsilon_{j,q}$ is a random variable supported in $[0,1]$. All $\varepsilon_{j,q}$ are independent, having the same law levelwise. For instance, $\varepsilon_{j,q}$ and $\varepsilon_{j,q'}$ are distributed identically with probability distribution p_j for all j, q, q'. In addition, we assume that $p_j(0) < 1$ for infinitely many j.

In line with the law p_j of the local regularity behavior of the functions in $\mathcal{S}$, we will consider the particular case of data with uniformly distributed wavelet coefficients. $F_g(\alpha) = 1$ for $\alpha = \gamma - 1/2$

$$\begin{aligned} \text{if} \quad & \left(\alpha = \gamma - \frac{1}{2}\right) F_g\left(\alpha\right) = 1 \\ \text{else} \quad & F_g\left(\alpha\right) = -\infty. \end{aligned} \tag{10}$$

Step 3. Our second type of stroke data has one of the simplest fractal stochastic processes, which is the fractional Brownian motion (fBm) (for more information see [29]). As is well known, fBm is the zero mean Gaussian process $X(t)$ with covariance function;

$$R\left(t,s\right) = \frac{\sigma^2}{2}\left(|t|^{2H} + |s|^{2H} - |t-s|^{2H}\right), \tag{11}$$

where H is a real number in $(0,1)$ and σ is a real number as well.

Step 4. Hence, the result of our denoising procedure will be wavelet-dependent in principle. The impact of the wavelet is controlled through the choice of the prior, that is, the multifractal, spectrum among all admissible ones. In practice, we have found out that few variations are observed if one uses a Daubechies wavelet [29] with length as 10 and a nonincreasing spectrum supported on $[H, \infty)$ with $F_g(H) = 1$.

The result of the denoised stroke dataset obtained by having applied the steps between Steps 1 and 4 on the stroke dataset is presented in Figure 2(a), mesh plot display for 2D mBd stroke dataset.

In this study, 2D mBd technique is applied to main captions of attributes (as can be seen in Table 2) pertaining to the stroke subtypes. 2D mBd by multifractal technique is reliant on the fact that stroke data enhancement is comparable to increasing the Hölder regularity at each point. Stroke dataset (X) was applied on FracLab [26] program to regularity-based enhancement $(\widehat{X})$ from 2D multifractal denoising techniques. Regularity-based enhancement of 2D mBd stroke dataset $\widehat{X}$ was clustered with K-means and FCM algorithms. Consequently, the most accurate clustering was attained for the subtypes of stroke.
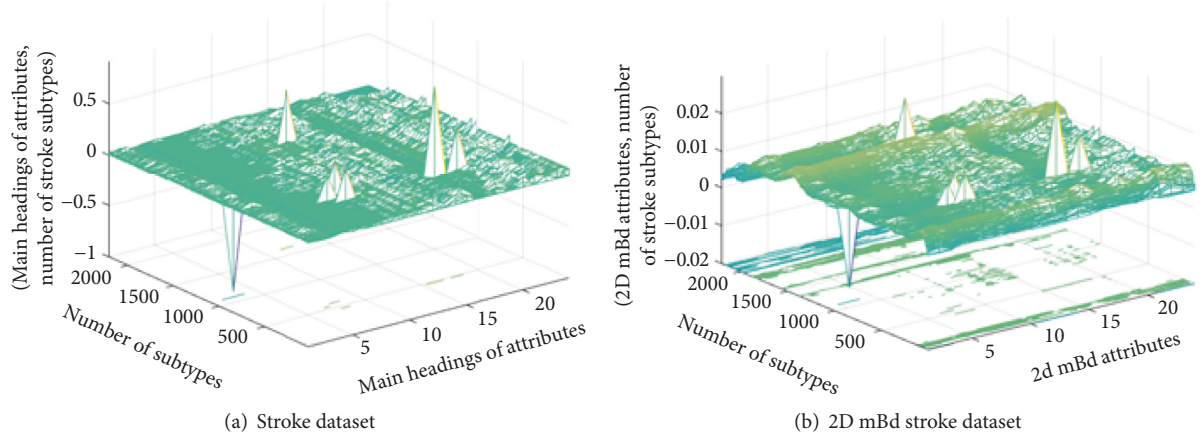
(a) Stroke dataset

(b) 2D mBd stroke dataset

FIGURE 2: Display of stroke dataset and 2D mBd stroke dataset with mesh plot.

(a) Stroke dataset

(b) Application of K-means and FCM algorithms

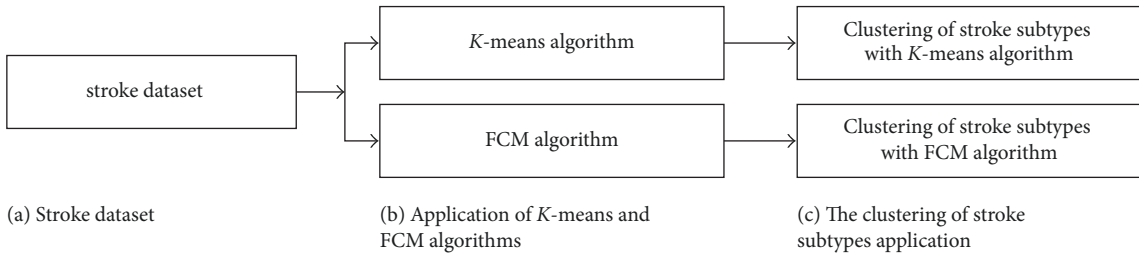(c) The clustering of stroke subtypes application

FIGURE 3: The application of the K-means and FCM algorithms to the stroke dataset.

2.2.4. K-Means Algorithm. $X = (x_1, x_2, \ldots, x_n)$ n d-dimensional data is to be grouped into a set of K clusters, $C = \{c_k, \ k = 1, \ldots, K\}$. K-means algorithm discovers a partition in which the squared error between the empirical mean of a cluster and the points in the cluster is reduced to minimum. With μ_k being the mean of cluster c_k, the squared error between μ_k and the points in cluster c_k can be defined based on the following equation [34–36]:

$$J(c_k) = \sum_{x_i \in c_k} \|x_i - \mu_k\|^2. \tag{12}$$

The aim of K-means is to minimize the sum of the squared error over all K clusters in line with

$$J(C) = \sum_{k=1}^{K} \sum_{x_i \in c_k} \|x_i - \mu_k\|^2. \tag{13}$$

K-means is initiated with a primary partition with K clusters and assigns patterns to clusters so that the squared error can be reduced. The squared error all the time goes down with an increase in the number of clusters K (with $J(C) = 0$); when $K = n$, it can be minimized only at a fixed number of clusters [37].

The clustering of the training set with the K-means clustering algorithm can be seen in Algorithm 1.

The key steps of K-means algorithm are as follows:

Step 1. Choose an initial partition with K clusters; repeat Steps 2 and 3 till the cluster membership stabilizes.

Steps 2-3. Form a new partition by assigning each pattern to its closest cluster center.

Step 4. Calculate the new cluster centers.

The clustering of the stroke dataset with the K-means clustering algorithm can be seen as follows:

Stroke dataset $X = (x_1, x_2, \ldots, x_{2204 \times 23})$ was applied to K-means algorithm. The main steps of K-means algorithm are as follows:

Step 1. K cluster is selected as an initial partition with 7 for $X = (x_1, x_2, \ldots, x_{2204 \times 23})$; Steps 2 and 3 are repeated up until the stroke subtypes cluster membership stabilizes.

Steps 2-3. For 1000 iterations, the data of each patient in the stroke dataset is assigned to the closest cluster centroid.

Step 4. The new cluster centroids of stroke subtypes are calculated. The results of the clustering procedure as obtained from the application of K-means algorithm (see Algorithm 1) on the stroke dataset are presented in Figure 4(a).

(a) Stroke dataset   (b) Multifractal denoising techniques   (c) Stroke dataset obtained from multifractal denoising techniques   (d) Application of K-means and FCM algorithms of the stroke dataset obtained from multifractal denoising techniques   (e) Clustering of stroke subtypes obtained from multifractal denoising techniques through K-means and FCM algorithms
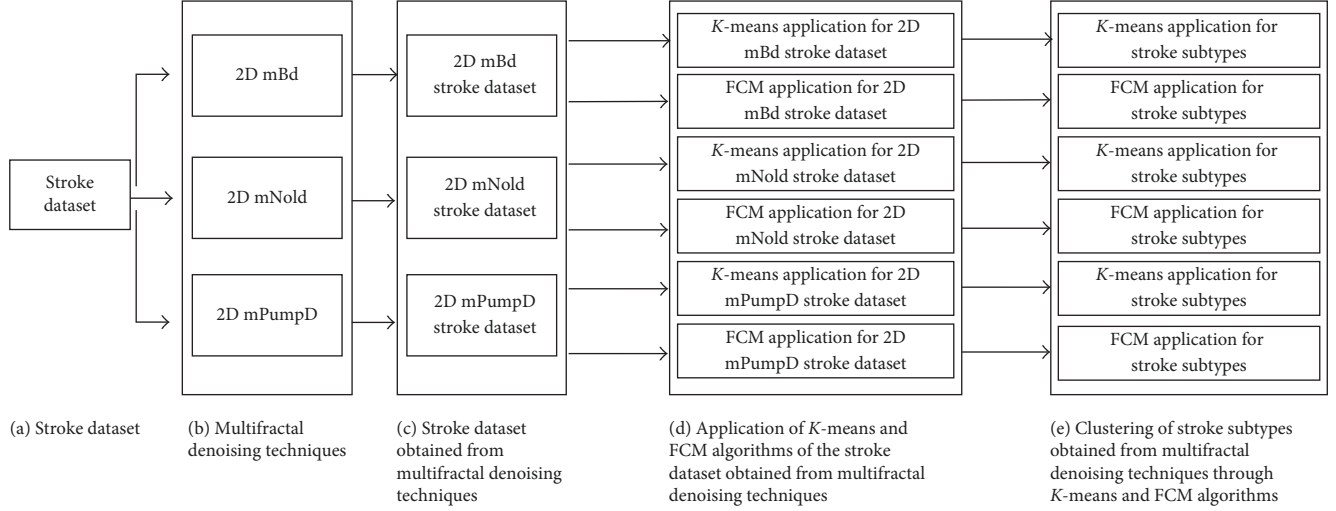
FIGURE 4: Clustering of the stroke dataset with the application of multifractal denoising techniques through K-means and FCM algorithms.

The clustering of the *2D mBd stroke dataset* with the K-means clustering algorithm can be described as follows:

The stroke datasets (mBd, mNold, mPumpD) obtained from the 2D multifractal denoising techniques $\widehat{X} = (\widehat{x}_1, \widehat{x}_2, \ldots, \widehat{x}_{2204 \times 23})$ were applied to K-means algorithm. The main steps of K-means algorithm are as follows:

Step 1. K cluster is selected as an initial partition with 7 for $\widehat{X} = (\widehat{x}_1, \widehat{x}_2, \ldots, \widehat{x}_{2204 \times 23})$; Steps 2 and 3 are repeated up until the stroke subtypes cluster membership stabilizes.

Steps 2-3. For 1000 iterations, the data of each patient in the stroke dataset is assigned to the closest cluster centroid.

Step 4. The new cluster centroids of stroke subtypes are calculated accordingly. The results of the clustering procedure as obtained from the application of K-means algorithm (see Algorithm 1) on the 2D mBd stroke dataset are presented in Figure 4(b).

2.2.5. Fuzzy C Means Algorithm. The FCM algorithm assigns data to each category through the use of fuzzy memberships [38–41]. Let $X = (x_1, x_2, \ldots, x_n)$, n-dimensional data signifying with N data to be split into c clusters, in which it denotes the features data. The iterative optimization algorithm steps are provided below:

$$J_m = \sum_{j=1}^{N} \sum_{i=1}^{C} u_{ji}^m \|x_j - c_i\|^2, \quad 1 \le m < \infty, \quad (14)$$

where u_{ji} denotes the membership of data x_j in the ith cluster, x_j is the jth cluster center, $\| \cdot \|$ is norm metric, and m is the denotation of a constant greater than 1.

The cost function is brought to the minimum when high membership values are assigned to the data that are close to the centroid of their clusters. In addition, the low membership values are assigned to the data which encompass data distant from the centroid. The membership function shows the probability that data belongs to a specific cluster.

As for FCM algorithm, the probability is dependent merely on the distance between the data and the cluster center of each stroke patient in the feature domain. The cluster centers and membership functions can be updated by using

$$u_{ji} = \frac{1}{\sum_{l=1}^{C} \left(\|x_j - c_i\| / \|x_j - c_l\| \right)^{2/(m-1)}};$$

$$1 \le j \le N, \ 1 \le i \le C, \quad (15)$$

$$c_i = \frac{\sum_{j=1}^{N} u_{ji}^m x_j}{\sum_{j=1}^{N} u_{ji}^m}.$$

The clustering of the training set with the FCM clustering algorithm is stated as in Algorithm 2.

Beginning with a preliminary guess for each cluster center, the FCM converges a solution for c_i, demonstrating the local minimum of the cost function. It is possible to detect such convergence through the comparison of the changes in the membership function or the cluster center at two consecutive iteration steps (l).

The clustering of the stroke dataset with the FCM clustering algorithm can be depicted as follows:

Stroke dataset $X = (x_1, x_2, \ldots, x_{2204 \times 23})$ was applied to the FCM algorithm. The main steps of FCM algorithm can be seen in the following steps regarding the stroke dataset:

Steps 1-3. $J = \sum_{j=1}^{2204 \times 23} \sum_{i=1}^{7} u_{ji}^2 \|x_j - c_i\|^2$ (based on (14)). Here, u_{ij} represents the membership of data x_j in the ith cluster, c_i is the ith cluster center, $\| \cdot \|$ is norm metric, and m is chosen as 2.

Step 4. The cost function is brought to the minimum when high membership values are assigned to the stroke dataset close to the cluster centroid for 1000 iterations. The results of the clustering procedure as obtained from the application of FCM algorithm (see Algorithm 2) on the stroke dataset are presented in Figure 5(a).
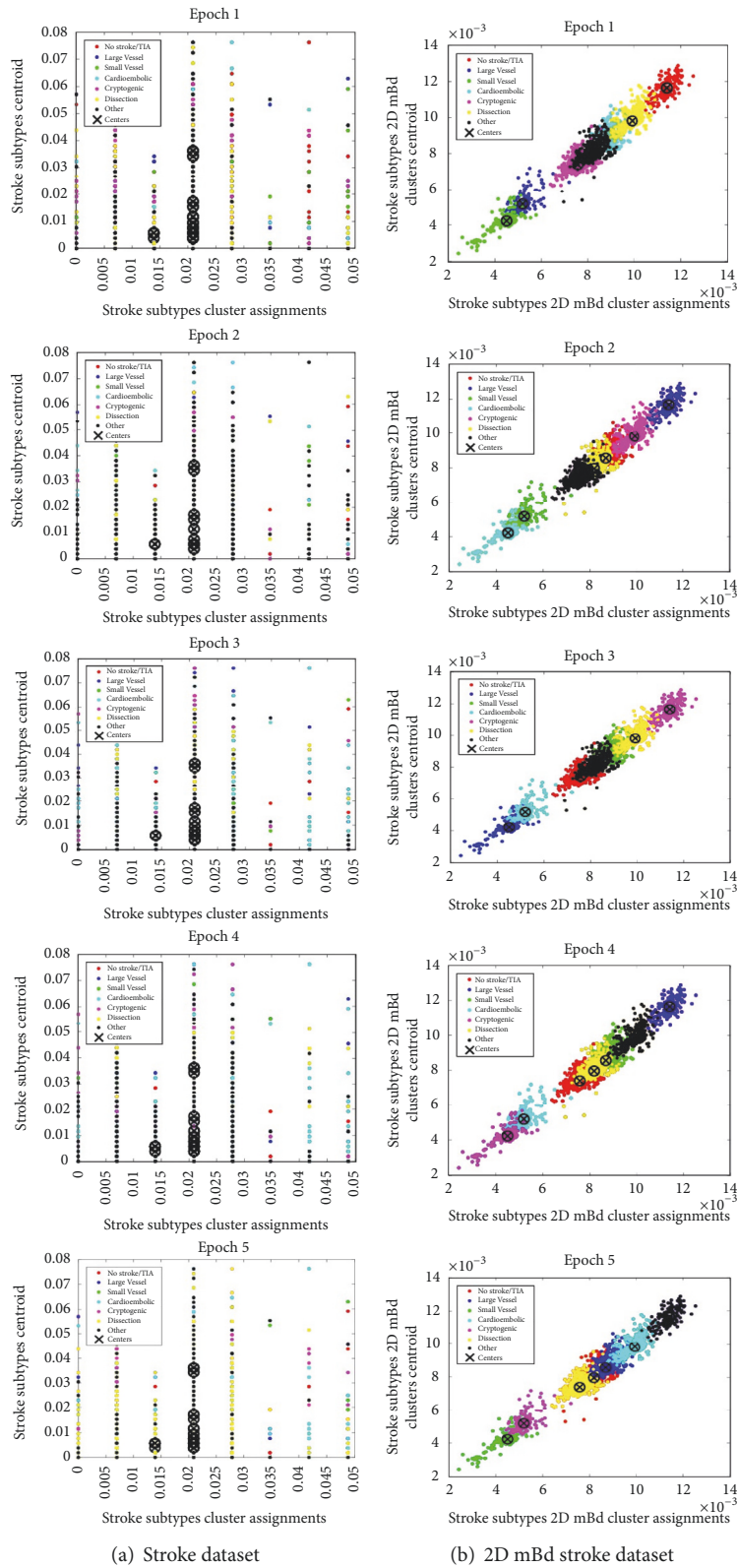
(a) Stroke dataset (b) 2D mBd stroke dataset

FIGURE 5: The clustering analyses of K-means algorithm based on the epochs.

Data: Training Set $X = (x_1, x_2, \ldots, x_n)$
Result: $J(C) = 0$
(1) Set $C < X$;
(2) **while** $J(C) = 0$ **do**
(3) $J(C) \longleftarrow \sum_{k=1}^{K} \sum_{x_i \in c_k} \|x_i - \mu_k\|^2$
(4)   find $J(c_k) \longleftarrow \sum_{x_i \in c_k} \|x_i - \mu_k\|^2$

ALGORITHM 1: K-means algorithm in the stroke dataset and 2D mBd stroke datasets.

Data: Training Set $X = (x_1, x_2, \ldots, x_n)$
Result: u_{ji}
(1) Set m, c;
(2) **while** c_i *local minimum* **do**
(3) $c_i \longleftarrow \dfrac{\sum_{j=1}^{N} u_{ji}^m x_j}{\sum_{j=1}^{N} u_{ji}^m}$
(4)   find $u_{ji} \longleftarrow 1/\sum_{l=1}^{C} (\|x_j - c_i\|/\|x_j - c_l\|)^{2/(m-1)}$

ALGORITHM 2: FCM algorithm in the stroke dataset and 2D mBd stroke datasets.

The clustering of the 2D mBd stroke dataset with the FCM clustering algorithm is as follows:

$\widehat{X} = (\widehat{x}_1, \widehat{x}_2, \ldots, \widehat{x}_{2204 \times 23})$ obtained from the *2D multifractal denoising techniques* were applied to the FCM algorithm. The main steps of FCM algorithm are stated below:

Step (1–3). $J = \sum_{j=1}^{2204 \times 23} \sum_{i=1}^{7} u_{ji}^2 \|\widehat{x}_j - c_i\|^2$ (based on (14)), where u_{ji} shows the membership of data $\widehat{x}_j$ in the ith cluster, c_i is the ith cluster centroid, $\|\cdot\|$ is norm metric, and m is chosen as 2.

Step (4). The cost function is minimized when high membership values are assigned to the stroke dataset close to the cluster centroids for 1000 iterations. The results of the clustering procedure as obtained from the application of FCM algorithm (see Algorithm 2) on the stroke dataset are presented in Figure 5(b).

3. Results and Discussion

In order to have a detailed vision of the relationship between the variables concerning the stroke dataset and 2D mBd stroke dataset in A of this study, the demonstration is done as plot based on mesh function. B includes the application of stroke dataset and regular and denoising 2D stroke dataset to the K-means algorithm, and C includes the calculations and results pertaining to the application of stroke dataset and regular and denoising 2D stroke dataset to the FCM algorithm.

3.1. Mesh Plot Display for Stroke Dataset and 2D mBd Stroke Dataset. The stroke dataset in our study is a matrix with a dimension of 2204×23. Figure 2(a) displays the meaning attribute headings for the stroke dataset attributes (demographic information, medical history, results of laboratory

TABLE 4: K-means algorithm parameters.

Parameters	Parameters value
K clusters	7
Maximum number of iterations	1000

tests, treatments, and medications) as well as the relationship between the stroke patients based on mesh function in plot.

Figure 2(a) presents the stroke dataset (X, Y, Z) drawing a wireframe mesh and a contour plot under it with color determined by Z. Thus, the color is proportional to the surface height. X and Y are vectors length (X: main headings of attributes) = 23 and length (Y = number of stroke subtypes) = 2204, where $[23, 2204]$ = size (Z = main headings of attributes, number of stroke subtypes). Here, $(X(j), Y(i), Z(i, j))$ are the wireframe grid lines' intersections; X and Y stand for the columns and row of Z.

Figure 2(a) displays the main headings of attributes concerned with the stroke dataset (demographic information, medical history, results of laboratory tests, treatments, and medications) and 2D mBd technique was applied to the data of stroke patients. Figure 2(b) shows the Bayesian regularity that matches each attribute in the stroke dataset (2204×23) as plot based on the mesh function result. Figure 2(b) presents the 2D mBd stroke dataset (X, Y, Z) drawing a wireframe mesh and a contour plot under it, with color determined by Z. Thus, the color is proportional to the surface height. X and Y are vectors length (X: 2D mBd attributes) = 23 and length (Y: 2D mBd number of stroke subtypes) = 2204, where $[23, 2204]$ = size (Z: 2D mBd attributes, 2D mBd number of stroke subtypes). Here, $(X(j), Y(i), Z(i, j))$ are the wireframe grid lines' intersections; X and Y stand for the columns and row of Z. In the stroke dataset of our study, Daubechies wavelet with a length ranging between 2 and 20 was applied. fBm with $H = 0.6$ and denoised version with Gaussian noise was applied on the dataset. As Figure 2(b) shows, taking f_g to be the theoretical spectrum is obtained with increments, or taking $f_g(\alpha) = 1$ for $\alpha \geq H$.

In this study, for the clustering procedure of stroke subtypes, K-means and FCM algorithms were applied to the stroke dataset (Figure 3(a)) as an initial step. The clustering procedure was obtained for the stroke subtypes (the details can be seen in Figure 3). In the second stage, multifractal denoising techniques (see Figure 4(b)) were applied to the same dataset (Figure 4(a)), stroke dataset. K-means and FCM algorithms were applied to the regular and denoised stroke dataset obtained (Figure 4(d)). As a result, the clustering procedure was obtained for the stroke subtypes (the details of which can be seen in Figure 4).

3.2. Application of K-Means Clustering Algorithm. Different iterations (200, 300, 400, 500, and 1000) have been used for the clustering procedure of stroke dataset, 2D mBd stroke datasets through the K-means algorithm. The most accurate results in the K-means algorithm have been obtained for the 1000 iterations. As shown in Figure 5, each epoch corresponds to 200 for more vivid display of classification.

The parameters pertaining to the 1000 iterations are presented in Table 4.

TABLE 5: Result of K-means algorithm clustering for 1000 iterations (as obtained from Figure 5).

Steps concerning epochs	$J(C)$ for stroke dataset (Figure 5(a))	$J(C)$ for 2D mBd stroke dataset (Figure 5(b))
Epoch 1 with 200 iterations	317.372	192.979
Epoch 2 with 200 iterations	319.855	19.2988
Epoch 3 with 200 iterations	320.911	192.984
Epoch 4 with 200 iterations	318.939	19.2983
Epoch 5 with 200 iterations	317.292	19.2979

In Figure 5(a) shows the best total sum of distance ($J(C)$) for the centroid values. The value ($J(C)$) did not change following this value. The clustering calculation was stopped in the 1000th iteration since no change was recorded in the results of the centroid values compared to the previous iteration.

In Figure 5(b), the result of best total sum of distance ($J(C)$) of the stroke 2D mBd dataset with K-means algorithm was obtained as 19.2979. The value ($J(C)$) did not change after this value. After the 900th iteration the iteration was stopped since no change in the iteration calculation happened in the clustering calculation compared to that of the previous iteration. Both of the datasets $J(C)$ have a 7-by-23 matrix that contains the final centroid locations. K-means is used to calculate the distance from each centroid to points on a grid. To be able to do this, the centroids $J(C)$ and points on a grid to K-means are passed, and 1000 iterations of the algorithm are implemented (Table 5).

The result of the best total sum of distance calculated (see Figure 5) reveals that 2D mBd stroke dataset has a better clustering accuracy than the stroke dataset.

The reason for mentioning such parameters is that these parameters were of help for the best cluster analysis to be performed in this study.

For the 1000 iterations in K-means clustering algorithm by splitting to 200 iterations with corresponding epoch, the calculation result pertaining to the stroke dataset is presented in Figure 5(a) for stroke dataset (see Figure 2(a)) and in Figure 5(b) for 2D mBd stroke dataset (see Figure 2(b)).

3.3. The Application of FCM Clustering Algorithm. Different iterations (200, 300, 400, 500, and 1000) have been used for the clustering procedure of stroke dataset, 2D mBd stroke datasets through the fuzzy C means algorithm. The most accurate results in the fuzzy C means algorithm have been obtained for the 1000 iterations. As shown in Figure 6, each epoch corresponds to 200 for more vivid display of classification.

For the 1000 iterations in FCM clustering algorithm, by splitting to 200 iterations with corresponding epoch, the calculation result pertaining to the stroke dataset is presented in Figure 6(a) for stroke dataset (see Figure 2(a)) and in Figure 6(b) for 2D mBd stroke dataset (see Figure 2(b)).

The parameters for the 1000 iterations are displayed in Table 6.

The reason for mentioning such parameters is that these parameters were of help for the best cluster analysis to be conducted in this study.

TABLE 6: FCM clustering algorithm parameters.

Parameters	Parameters value
Exponent for the partition matrix U	2.0
Maximum number of iterations	200
Minimum amount of improvement	$1e - 3$

The data in stroke dataset (2204×23) and 2D mBd stroke dataset (2204×23) as well as the distance between the cluster centroids and the data are computed (the distance between the data and the cluster center of each stroke patient in the feature domain). In addition, the results computed are stored in the u_{ji} matrix. The optimization stopped on the 1000th iteration for the stroke dataset (in Figure 6(a)) and 2D mBd stroke dataset (in Figure 6(b)), as the objective function improved by less than $1e - 3$ between the final two iterations. The clustering process comes to an end when the maximum number of iterations is reached, or when the objective function improvement between two consecutive iterations is less than the specified minimum. In the calculations of u_{ji} matrix for 1000 iterations (see Figure 6), it has been revealed that the clustering accuracy of 2D mBd stroke dataset is better than that of the stroke dataset (Table 7).

Calculations given in Table 7 were performed using Matlab, Mosek, and FracLab [26] environment. The clustering accuracy rates obtained from K-means and FCM algorithms applied to the stroke dataset and 2D mBd, 2D mNold, 2D mPumpD stroke datasets in the study are provided in Table 8.

Accurate clustering results have been obtained with K-means and FCM algorithms by applying 2D multifractal denoising techniques for the stroke dataset based on the clustering results of the stroke subtypes obtained in this study (Table 8). The clustering accuracy of the stroke subtypes with K-means and FCM algorithms through the 2D mBd stroke dataset proved to be better compared to the 2D mNold stroke dataset and 2D mPumpD stroke dataset. In the 2D mBd stroke dataset, the clustering accuracy of the K-means algorithm proves to be better for the no stroke/TIA and cardioembolic stroke subtypes. In the 2D mBd stroke dataset, the clustering results regarding the FCM algorithm application proved to be more accurate for the large vessel, small vessel, cryptogenic, dissection, and other stroke subtypes.

In literature, very limited number of papers exist on mathematical modeling and clustering of stroke subtypes. In this study, seven subtypes among no stroke/TIA, large vessel, small vessel, cardioembolic, cryptogenic, dissection, and other subtypes have been analyzed. Four datasets (stroke

(a) Stroke dataset   (b) 2D mBd stroke dataset
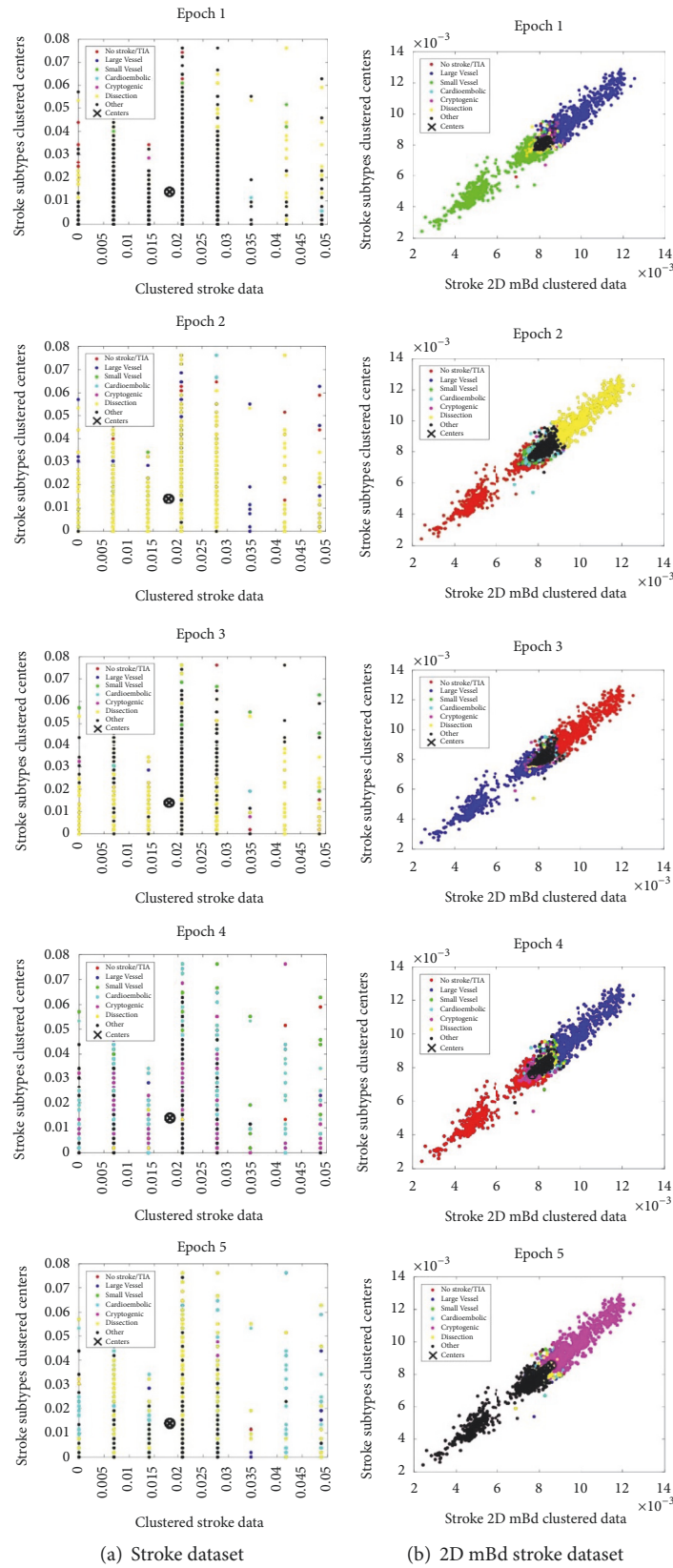
Figure 6: The clustering analyses of FCM algorithm based on the epochs.

TABLE 7: Result of FCM algorithm clustering with for 1000 Iterations (as obtained from Figure 6).

Steps concerning epochs	Objective function vector for stroke dataset (Figure 6(a))	Objective function vector for 2D mBd stroke dataset (Figure 6(b))
Epoch 1 with 200 iterations	[3.2285, 2.4407, 2.4406]	[0.0318, 0.239, 0.237]
Epoch 2 with 200 iterations	[3.1957, 2.4407, 2.4406]	[0.0319, 0.239, 0.237]
Epoch 3 with 200 iterations	[3.2490, 2.4407, 2.4406]	[0.0317, 0.238, 0.235]
Epoch 4 with 200 iterations	[3.2824, 2.4407, 2.4406]	[0.0316, 0.238, 0.234]
Epoch 5 with 200 iterations	[3.2255, 2.4407, 2.4406]	[0.0316, 0.239, 0.238]

TABLE 8: Clustering results.

Stroke Subtypes	K-means stroke data (%)	K-means 2D mBd stroke data (%)	K-means 2D mNold stroke data (%)	K-means 2D mPumpD stroke data (%)	FCM stroke data (%)	FCM 2D mBd stroke data (%)	FCM 2D mNold stroke data (%)	FCM 2D mPumpD stroke data (%)
No stroke/TIA	34.7	83	52.1	46.4	30	64	39	49
Large vessel	58.2	70.7	60.2	68.6	8.3	83.1	62	83
Small vessel	63.3	76	71.3	71.2	17	96.5	87	76
Cardioembolic	17.4	69.6	34.5	43.5	38	63	43	60
Cryptogenic	2.4	62.5	49.2	56.8	15	98	76.6	41
Dissection	19.6	47.1	17.3	17.3	10	67.7	26.8	14
Others	14.8	20.1	16.4	12.4	12	62.4	26	15.6

dataset, 2D mBd stroke dataset, 2D mNold stroke dataset, 2D mPumpD stroke dataset) are totally performed on our new approach with K-means and FCM algorithms.

4. Conclusions

The main contribution of this paper is that it has proposed a novel approach in stroke main headings regarding attributes with the use of 2D techniques of multifractal denoising. The clustering performances of 2D multifractal denoised techniques (mBd, mNold, mPumpD) for the stroke subtypes regarding a total of 2204 stroke patients' dataset have been provided in a comparative manner. When our study is compared with the other works [22, 23], it is seen that first of all there is no attribute constraint for the clustering of 7 subtypes of stroke. Secondly, it is possible to select the efficient and significant attributes through the denoising techniques. Finally, popular FCM and K-means algorithms are applied to the datasets comprised of efficient and significant attributes. The output datasets are provided in supervised learning and they are used to train the machine and get the desired outputs, whereas in unsupervised learning, no datasets are provided. Instead, the data is clustered into different classes. K-means is one of the simplest unsupervised learning algorithms that solve the well-known clustering problem. The procedure follows a simple way to classify a given dataset through a certain number of clusters (assuming K clusters) fixed a priori. K-means is a simple algorithm that has been adapted to many problem domains. As we have seen, it is a good candidate for extension to work with fuzzy feature vectors [42]. Fuzzy C means algorithm [43, 44] yields the best result for overlapped dataset and is comparatively better than K-means algorithm. Unlike K-means, where data point must exclusively belong to one cluster center, here the data point

is assigned membership to each cluster center as a result of which data point may belong to more than one cluster center. For the first time in literature, 2D multifractal denoising techniques and K-means and FCM algorithms have been applied to the numeric data obtained from the attributes that belong to the patients with seven different stroke subtypes. The results reveal that the 2D Bayesian denoising technique application used in our study has proven to be much better compared to the other techniques and methods.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

References

[1] B. Michael, *Fractals Everywhere*, Academic Press, San Diego, Calif, USA, 1988.

[2] D. Baleanu, J. A. Tenreiro Machado, C. Cattani, M. C. Baleanu, and X.-J. Yang, "Local fractional variational iteration and decomposition methods for wave equation on cantor sets within local fractional operators," *Abstract and Applied Analysis*, vol. 2014, Article ID 535048, 6 pages, 2014.

[3] A. L. Goldberger, L. A. N. Amaral, J. M. Hausdorff, P. C. Ivanov, C.-K. Peng, and H. E. Stanley, "Fractal dynamics in physiology: alterations with disease and aging," *Proceedings of the National Acadamy of Sciences of the United States of America*, vol. 99, no. 1, pp. 2466–2472, 2002.

[4] J. Helenius, R. P. Goddeau, M. Moonis, and N. Henninger, "Impact of leukoaraiosis burden on hemispheric lateralization

of the national institutes of health stroke scale deficit in acute ischemic stroke," *Stroke*, vol. 47, no. 1, pp. 24–30, 2016.

[5] Y. Mayasi, J. Helenius, D. D. McManus et al., "Atrial fibrillation is associated with anterior predominant white matter lesions in patients presenting with embolic stroke," *Journal of Neurology, Neurosurgery & Psychiatry*, vol. 47, no. 30, pp. 24–30, 2017.

[6] A. Minaeian, A. Patel, B. Essa, R. P. Goddeau, M. Moonis, and N. Henninger, "Emergency department length of stay and outcome after ischemic stroke," *Journal of Stroke and Cerebrovascular Diseases*, vol. 26, no. 10, pp. 2167–2173, 2017.

[7] J. Adamson, A. Beswick, and S. Ebrahim, "Is stroke the most common cause of disability?" *Journal of Stroke and Cerebrovascular Diseases*, vol. 13, no. 4, pp. 171–177, 2004.

[8] H. Zhao, L. Pesavento, S. Coote et al., "Paramedic validation of an australian large vessel occlusion triage algorithm for stroke," *Journal of Neurology, Neurosurgery & Psychiatry*, vol. 88, no. 5, 2017.

[9] L. Pantoni, "Cerebral small vessel disease: from pathogenesis and clinical characteristics to therapeutic challenges," *The Lancet Neurology*, vol. 9, no. 7, pp. 689–701, 2010.

[10] Z. Zhang, L.-J. Liu, C. Zhang, and Y.-P. Yu, "Association between interleukin-1 gene single nucleotide polymorphisms and ischemic stroke classified by toast criteria in the Han population of Northern China," *BioMed Research International*, vol. 2013, Article ID 961039, 6 pages, 2013.

[11] E. Kim, P. Jones, C. T. Richards et al., "Abstract WP272: impact of stroke legislation on developing stroke systems of care and improving acute therapy," *Stroke*, vol. 48, no. 1, 2017.

[12] S. Wang, M. Yang, Y. Zhang et al., "Detection of left-sided and right-sided hearing loss via fractional Fourier transform," *Entropy*, vol. 18, no. 5, article 194, 10 pages, 2016.

[13] G. Yang, Y. Zhang, J. Yang et al., "Automated classification of brain images using wavelet-energy and biogeography-based optimization," *Multimedia Tools and Applications*, vol. 75, no. 23, pp. 15601–15617, 2016.

[14] Y. Karaca and C. Cattani, "Clustering multiple sclerosis subgroups with multifractal methods and self-organizing map algorithm," *Fractals*, vol. 25, no. 4, Article ID 1740001, 10 pages, 2017.

[15] G. G. Tsaneva, "Denoising of electrocardiogram data with methods of wavelet transform," in *Proceedings of the International Conference on Computer Systems and Technologies*, June 2013.

[16] F. N. Doubal, T. J. MacGillivray, N. Patton, B. Dhillon, M. S. Dennis, and J. M. Wardlaw, "Fractal analysis of retinal vessels suggests that a distinct vasculopathy causes lacunar stroke," *Neurology*, vol. 74, no. 14, pp. 1102–1107, 2010.

[17] P. Shanmugavadivu, V. Sivakumar, and R. Sudhir, "Fractal dimension-bound spatio-temporal analysis of digital mammograms," *The European Physical Journal Special Topics*, vol. 225, no. 1, pp. 137–146, 2016.

[18] H. Ahammer, T. T. J. Devaney, and H. A. Tritthart, "Fractal dimension for a cancer invasion model," *Fractals*, vol. 9, no. 1, pp. 61–76, 2001.

[19] C. M. Helgason, D. S. Malik, S.-C. Cheng, T. H. Jobe, and J. N. Mordeson, "Statistical versus fuzzy measures of variable interaction in patients with stroke," *Neuroepidemiology*, vol. 20, no. 2, pp. 77–84, 2001.

[20] M. Rickhag, T. Wieloch, G. Gidö et al., "Comprehensive regional and temporal gene expression profiling of the rat brain during the first 24 h after experimental stroke identifies dynamic ischemia-induced gene expression patterns, and reveals a biphasic activation of genes in surviving tissue," *Journal of Neurochemistry*, vol. 96, no. 1, pp. 14–29, 2006.

[21] M. D. Stricker, N. C. Onland-Moret, J. M. A. Boer et al., "Dietary patterns derived from principal component- and k-means cluster analysis: Long-term association with coronary heart disease and stroke," *Nutrition, Metabolism & Cardiovascular Diseases*, vol. 23, no. 3, pp. 250–256, 2013.

[22] A. N. Pavlov, A. S. Abdurashitov, O. A. Sindeeva et al., "Characterizing cerebrovascular dynamics with the wavelet-based multifractal formalism," *Physica A*, vol. 442, pp. 149–155, 2016.

[23] S. Keerthana and K. Sathiyakumari, "Brain stroke segmentation using fuzzy C-means clustering," *International Journal of Computer Applications*, vol. 154, no. 4, pp. 26–30, 2016.

[24] H.-G. Bu, J. Wang, and X.-B. Huang, "Fabric defect detection based on multiple fractal features and support vector data description," *Engineering Applications of Artificial Intelligence*, vol. 22, no. 2, pp. 224–235, 2009.

[25] Y.-T. Chen, E. W. Sun, and M.-T. Yu, "Improving model performance with the integrated wavelet denoising method," *Studies in Nonlinear Dynamics and Econometrics*, vol. 19, no. 4, pp. 445–467, 2015.

[26] J. L. Vehel, FracLab, 2000, https://project.inria.fr/fraclab/.

[27] M. F. Barnsley, D. Saupe, and E. R. Vrscay, "Signal enhancement based on Holder regularity analysis," *IMA Volumes in Mathematics and Its Applications*, vol. 132, pp. 197–209, 2002.

[28] Y. Meyer, *Wavelets, Vibrations and Scalings*, vol. 9 of *CRM Monograph Series*, American Mathematical Society, Providence, RI, USA, 1997.

[29] J. Lévy Véhel and P. Legrand, "Bayesian multifractal signal denoising," in *Proceedings of the IEEE International Conference on Accoustics, Speech, and Signal Processing*, pp. 177–180, April 2003.

[30] J. L. Véhel, "Numerical computation of the large deviation multifractal spectrum," in *CFIC'96—Chaos and Fractals in Chemical Engineering*, pp. 1–3, World Scientific Publishing, 1996.

[31] M. Li, Z. Wang, J. Luo, Y. Liu, and S. Cai, "Wavelet denoising of vehicle platform vibration signal based on threshold neural network," *Shock and Vibration*, vol. 2017, Article ID 7962828, 12 pages, 2017.

[32] O. Nicolis, P. Ramirez-Cobo, and B. Vidakovic, "2D wavelet-based spectra with applications," *Computational Statistics & Data Analysis*, vol. 55, no. 1, pp. 738–751, 2011.

[33] E. Lutton, P. Grenier, and J. L. Vehel, "An interactive EA for multifractal bayesian denoising," in *EvoWorkshops 2005: Applications of Evolutionary Computing*, pp. 274–283, Springer, 2005.

[34] A. K. Jain, "Data clustering: 50 years beyond K-means," *Pattern Recognition Letters*, vol. 31, no. 8, pp. 651–666, 2010.

[35] T. Kanungo, D. M. Mount, N. S. Netanyahu, C. D. Piatko, R. Silverman, and A. Y. Wu, "An efficient k-means clustering algorithms: analysis and implementation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 7, pp. 881–892, 2002.

[36] P. K. Agarwal and C. M. Procopiuc, "Exact and approximation algorithms for clustering," in *Proceedings of the 9th Annual ACM-SIAM Symposium on Discrete Algorithms*, vol. 33, no. 2, pp. 201–226, June 1998.

[37] K. Alsabti, S. Ranka, and V. Singh, "An efficient K-means clustering algorithm," in *Proceedings of the 1st Workshop High Performance Data Mining*, 1998.

[38] J. C. Bezdek, "FCM: the fuzzy c-means clustering algorithm," *Computers & Geosciences*, vol. 10, no. 2-3, pp. 191–203, 2003.

[39] N. R. Pal and J. C. Bezdek, "On cluster validity for the fuzzy c-means model," *IEEE Transactions on Fuzzy Systems*, vol. 3, no. 3, pp. 370–379, 1995.

[40] B. Nadler, S. Lafon, R. R. Coifman, and I. G. Kevrekidis, "Diffusion maps, spectral clustering and reaction coordinates of dynamical systems," *Applied and Computational Harmonic Analysis* , vol. 21, no. 1, pp. 113–127, 2006.

[41] W. Wu, S. Wu, Z. Zhou, R. Zhang, and Y. Zhang, "3D liver tumor segmentation in CT images using improved Fuzzy C-Means and graph cuts," *BioMed Research International*, vol. 2017, Article ID 5207685, 11 pages, 2017.

[42] J. MacQueen, "Some methods for classification and analysis of multivariate observations," in *Proceedings of the 5th Berkeley Symposium on Mathematical Statistics and Probability*, vol. 1, no. 14, pp. 281–297, June 1967.

[43] J. C. Bezdek, *Fuzzy mathematics in pattern classification [PhD thesis]*, Cornell University, Ithaca, NY, USA, 1973.

[44] J. C. Bezdek, *Pattern Recognition with Fuzzy Objective Function Algorithms*, Kluwer Academic Publishers, Norwell, Mass, USA, 1981.