Comparative analysis of machine learning classification of time series with fractal properties

Tamara Radivilova
dept. of Infocommunication Engineering
Kharkiv National University of Radio Electronics
Kharkiv, Ukraine
tamara.radivilova@gmail.com

Lyudmyla Kirichenko, Bulakh Vitalii
dept. of Applying mathematics
Kharkiv National University of Radio Electronics
Kharkiv, Ukraine
lyudmyla.kirichenko@nure.ua, bulakhvitalii@gmail.com

***Abstract*— The article analyses the classification of time series according to their fractal properties by machine learning. The classification was carried out using neural networks and the random forest method. Objects were the model fractal time series with given the Hurst exponent. Each class was a set of time series with the Hurst exponent values in a predetermined range. Input features were the values of time series. It was demonstrated that in this case the classification accuracy is high enough. The most accurate classification results were obtained using recurrent neural network. The proposed method can be readily used in practice for recognition, classification and clustering of time series with fractal properties.**

Keywords*— *fractal time series, time series classification, Hurst exponent, random forest , neural networks

I. Introduction

Optoelectronic and laser devices are used in many areas of human life, for example, in medicine, remote sensing, reflectometry, in optical communication networks, etc. Many data in optoelectronic devices can be represented as time series, for example, the intensity of continuous laser radiation, laser behavior, speckle fields and structures, and many others. All these data are high-speed signals depending on time and their processing can turn into an extremely difficult task, especially when signals nonlinearly distorted. A Machine Learning methods are used for fast processing of information, for example, when solving a problem of classifying data of highly distorted optical communication signals after an extended transmission over fiber [1, 2].

As studies have shown in recent decades, many complex technical systems and processes have fractal properties. These are infocommunication networks and information transmission devices, antenna and radar complexes, wave scattering processes, the turbulent motion of fluid and plasma, and much more [3, 4]. Among them, can be specified typical processes for optoelectronic and laser engineering. For example, a generation of fractal light inside a laser cavity [5] was recently demonstrated and in paper [6] fractal graphene layers are obtained by a complex method of liquid phase exfoliation and self-organization.

A number of applications of fractal geometry and fractal analysis in the field of laser optics are given in [2, 7-9]. The dynamics of such systems are represented by signals with fractal properties. In many cases, there are tasks of classification [1, 10, 11], including classification of fractal time series. Most often this happens by evaluating and analyzing fractal characteristics [12-18]. However, in recent years, methods of machine learning have been used to analyze and classify fractal series [19, 20]. This approach is fully applicable to the analysis of signals from optoelectronic devices [9].

In the above-mentioned works, it is shown that a qualitative change in the state of a complex dynamic system leads to a change in the fractal properties of its signals. Thus, the task of classifying time series on the basis of their fractal properties by machine learning methods is relevant in the field of optoelectronics and laser engineering.

The results of studies of fractal time series indicate that changes in quantitative fractal characteristics correspond to changes in the correlation and spectral structure of the series. In turn, a change in correlation leads to a change in the values of the time series. Thus, it can be assumed that the fractal properties of a time series can be classified on the basis of the values of the series itself. In the works [19-21] the possibility of classifying time series on the basis of fractal properties with the help of ensembles of decision trees was demonstrated.

The aim of this work is a comparative analysis of the classification of fractal time series by the methods of decision trees and neural networks using the values of time series as features.

II. Input Data: Fractal Time Series

The self-similarity of random processes is preservation of the distribution laws when changing the time scale, i.e.:

$$\mathrm{Law}\{X(t)\} = \mathrm{Law}\{a^{-H} X(at)\} \ \ \forall a > 0,\, t > 0 \ . \qquad (1)$$

It should be noted that if the process is represented by a time series $X = (X_1, X_2, \ldots)$, then to define self-similarity, the concept of an aggregate series is used. Let's denote by $X^{(m)} = \{X_1^{(m)}, X_2^{(m)}, \ldots)$ the averaged over blocks of length m process X, whose components are determined by the equality

$$X_t^{(m)} = \frac{1}{m}(X_{tm-m+1} + ... + X_{tm}), \;\; m,t \in N. \quad (2)$$

The time series has self-similarity with the parameter H , if it satisfies equation

$$\text{Law}\left\{m^{1-H} X^{(m)}\right\} = \text{Law}\left\{X\right\}, \quad (3)$$

that is, the process preserves the distribution laws after averaging over blocks.

The parameter of self-similarity $H,\ 0 < H < 1$ is called the Hurst exponent. Besides, parameter H characterizes the measure of long-term dependence of the process. In the case $0.5 < H < 1$ the process has a long memory: if for some time in the past a positive process increments were observed, i.e. there was an increase, then it will continue to increase the average. In the case $0 < H < 0.5$ high process values follow low, and vice versa. With $H = 0.5$ the deviations of the process from the average are really random and do not depend on the previous values.

Multifractal random processes are statistically heterogeneous fractal processes and have more flexible scaling.

$$\text{Law}\{X(at)\} = \text{Law}\{M(a) \cdot X(t)\}, \quad (4)$$

where $M(a)$ is an independent of $X(t)$ random function with specific properties. In the case of a self-similar process $M(a) = a^H$.

The degree of self-similarity and the long-term dependence of multifractal processes is also determined by the Hurst exponent H .

One of the most well-known and simple models of stochastic dynamics that have fractal properties is the fractal Brownian motion (fBm), which is widely used in physics, chemistry, biology, economics, and the theory of network traffic.

A random process $X(t)$ is called a fractal Brownian motion with a parameter $H,\ 0 < H < 1$, if its increments $\Delta X(\tau) = X(t+\tau) - X(t)$ have a normal distribution with zero expectation and variance $\text{Var}[X(t+\tau) - X(t)] = \sigma_0{}^2 \tau^{2H}$, where σ_0 is the diffusion coefficient.

The fBm increments are called fractal Gaussian noise (fGn). There are several methods for constructing fBmM for the case of discrete time [3]. A simple model of a multifractal process using fGn is presented in [22]. In this case, a multifractal time series is obtained on the basis of an exponential transform

$$Y(t) = \text{Exp}[k * X(t)], \quad (5)$$

where $X(t)$ is the time series of fGn with a given Hurst exponent H , k is some coefficient affecting the degree of multifractality time series $Y(t)$. Using transform (5), multifractal realizations are obtained with given fractal properties, in particular, with a given Hurst exponent

III. Classification Methods

1) The random forest methods

A decision tree is a decision support tool used in data analysis for models that predict the value of the target variable based on several input variables. The method of decision trees is considered one of the simplest and most effective for solving classification problems, arising in many different areas. It consists in splitting the initial data into groups according to some splitting rules until their homogeneous subsets are obtained. One of the disadvantages of the method is that the decision tree changes quite a lot with a relatively small change in the data sample.

Random forest is an ensemble of decision trees and can be considered as a complex, composite model [23]. When combining several trees into a composition scatter of target variable values may become substantially less. The important point is the element of randomness in the creation of each tree and the random selection of features. The work [20] showed that the best results of time series classification were shown by the random forest algorithm based on regression trees. When using regression trees, the result of the model is the probability of matching the time series to a given class.

Thus, in this paper, the random forest algorithm using regression trees was chosen as one of the methods for classifying fractal time series.

2) Neural networks

To solve the problem of time series classification using neural networks, two different neural network architectures were selected and investigated during the experiment.

The first neural network was made up of seven fully connected layers of large dimension with the activation function of the ReLU type. The use of this activation function requires much fewer resource costs and significantly increases the rate of convergence of stochastic gradient descent compared with the popular sigmoid and hyperbolic tangent activation functions. After each full layer, the regularization layer was included in the network. The method of regularization was chosen batch normalization [24], which is used to stabilize the neural network, to prevent the effect of overtraining and increase productivity.

The second network had one recurrent layer for taking into account the relationship between elements and six fully connected layers. In this case, a regularization layer was also included after each fully connected layer. The stochastic optimization method Adam (Adaptive Moment Estimation) [25], which is an extension method of stochastic gradient descent, was chosen as a method of learning both neural networks.

IV. Experiment Description and Results

In this paper, each class was a set of time series generated using expression (5) with the Hurst exponent belonging to a given range of values for this class. For each time series, the Hurst exponent was chosen randomly within the appropriate range. The values of the Hurst exponent are changing in the range from 0.5 to 1 with a step of 0.1. The minimum and maximum values of the Hurst exponent were selected 0.51 and 0.99, respectively. Thus, the training of models was carried out in 5 classes, where H∈{[0.51, 0.6), [0.6, 0.7), [0.7, 0.8), [0.8, 0.9] , [0.8, 0.99]}.

Figure 1 shows typical time series from different classes: at the top is a series with the Hurst exponent *H=0.75*, at the bottom a series for which *H=0.85*.

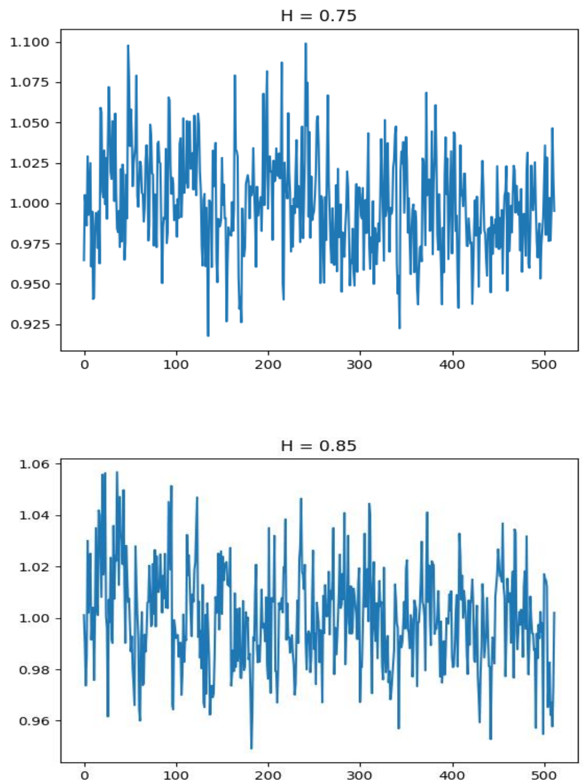

Fig. 1. Classified time series from different classes

For the experiment, three models were used: regression decision trees with random forest, neural network, recurrent neural network. At the input of the models were submitted data of time series without additional transformations, i.e. features of the models were the values of the time series itself. The result of each model was the probability of matching the time series to each class.

To build classifiers and conducting research, a high-level programming language Python was chosen, which contains libraries that implement a variety of machine learning methods, including neural networks. The training of models for each class was carried out on 300 examples of time series and tested for 50 test cases.

The classification was carried out for time series of different lengths, however, to compare the results the focus was on series with a length of 500 values. On the one hand, such a length of time series is typical for a range of values of many real time series, on the other hand, for such a time series length, methods for estimating quantitative fractal characteristics have a large error.

In fig. 2 shows the histograms of the probability distribution of the class number prediction for each range of values of the Hurst exponent, obtained using a neural network. Above each histogram is shown the value of *H*, equal to the middle of the range. Such distributions are typical for all variants of classification.

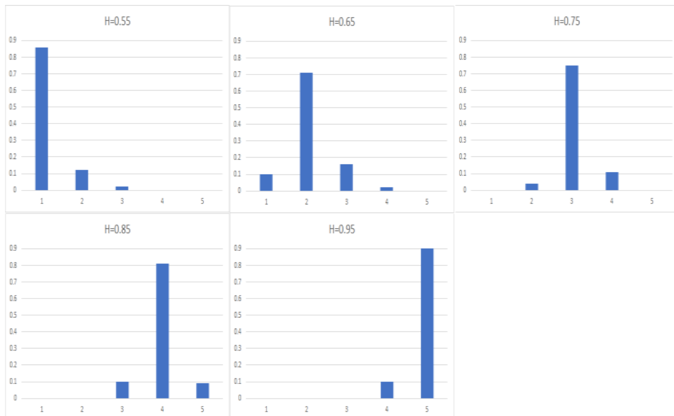

Fig. 2. The probability distribution for determining the class number depending on the range of the Hurst exponent for a neural network

Table 1 presents the average probability of determining the class for the three classifiers: neural network with ReLU (NN), a neural network with a recurrent layer (RNN) and Random Forest (RF), depending on the range of values of the Hurst exponent. The last column contains the probabilities averaged over all classes.

TABLE I. THE AVERAGE PROBABILITY OF CLASS PREDICTION FOR DIFFERENT CLASSIFIERS

Classifiers	Ranges of the Hurst exponent					
	[0.51,0.6)	[0.6,0.7)	[0.7,0.8)	[0.8,0.9)	[0.9,0.99]	
RF	0.75	0.78	0.82	0.81	0.76	0.78
NN	0.86	0.71	0.75	0.81	0.73	0.77
RNN	0.85	0.72	0.73	0.88	0.9	0.82

Thus, the results indicate the possibility of classifying fractal time series by machine learning methods with high accuracy. The classification results presented in the table demonstrate some advantage of the recurrent neural network over the neural network with ReLU and the ensemble of regression decision trees.

V. Conclusion

A comparative analysis of the machine learning classification of stochastic time series based on their fractal properties has been carried out. The results have shown that for detecting the degree of time series self-similarity, determined by the Hurst exponent, can be successfully used the time series values as features.

Neural networks with various architecture and Random Forest algorithm have been used as classifiers. The neural network with a recurrent layer showed slightly better results in classification accuracy. The method of using decision trees is not inferior in the accuracy of the neural network.

In our future research, we intend to concentrate on improvement of the recurrent neural network architecture for the classification of real fractal series.

References

[1] J. H. García-López, R. Jaimes-Reátegui, S. M. Afanador-Delgado, R. Sevilla-Escoboza, G. Huerta-Cuéllar, D. López-Mancilla, R. Chiu-Zarate, C. E. Castañeda-Hernández and A. Pisarchik, "Experimental and Numerical Study of an Optoelectronics Flexible Logic Gate Using a Chaotic Doped Fiber Laser." *Recent Development in Optoelectronic Devices*, Chapter 6, pp.97-114, 2018. doi: http://dx.doi.org/10.5772/intechopen.75466

[2] Tatiana E. Itina (Ed.), "*Laser Ablation - From Fundamentals to Applications*." Intech, 2017. doi: 10.5772/67961

[3] O. I. Shelukhin, S. M. Smolskiy, A. V. Osin, "*Self-Similar Processes in Telecommunications*," John Wiley & Sons, New York, pp. 183-196 (2007). doi: https://doi.org/10.1002/9780470062098

[4] F. Brambila (Ed.), "*Fractal Analysis - Applications in Physics, Engineering and Technology*." Intech, 2017. doi: 10.5772/65531

[5] H. Sroor, D. Naidoo, S. W. Miller, J. Nelson, J. Courtial and A. Forbes, "Fractal light from lasers," *Physical Review A* 99, 013848, 2019. doi: 10.1103/PhysRevA.99.013848

[6] Narek Margaryan, Ninel Kokanyan, Edvard Kokanyan, "Low-temperature synthesis and characteristics of fractal graphene layers." *Journal of Saudi Chemical Society*, vol. 23, issue 1, pp. 13-20, 2019. doi: https://doi.org/10.1016/j.jscs.2018.03.004

[7] J. M. Christian, G. S. McDonald, A. S. Heyes and J. G. Huang, "Kaleidoscope laser properties and new optical fractal contexts," *2010 International Conference on Advanced Optoelectronics and Lasers*, Sevastopol, 2010, pp. 130-132. doi: 10.1109/CAOL.2010.5634238

[8] V. M. Shalaev, "*Nonlinear Optics of Random Media: Fractal Composites and Metal-Dielectric Films*." Springer, 2007.

[9] S. Phuvan, T. K. Oh, N. Caviris, Y. Li and H. Szu, "Optoelectronic fractal scanning technique for wavelet transform and neural net pattern classifiers," [*Proceedings 1992*] *IJCNN International Joint Conference on Neural Networks*, Baltimore, MD, USA, 1992, pp. 40-46 vol.1. doi: 10.1109/IJCNN.1992.287240

[10] J. Matuszewski, "Radar signal identification using a neural network and pattern recognition methods," 2018 14th International Conference on Advanced Trends in Radioelecrtronics, Telecommunications and Computer Engineering (TCSET), Lviv-Slavske, 2018, pp. 79-83. doi: 10.1109/TCSET.2018.8336160

[11] V. P. Mashtalir, V. V. Shlyakhov and S. V. Yakovlev, "Group structures on quotient sets in classification problems." *Cybernetics and Systems Analysis,* vol.50. №4. pp. 507-518.

[12] L. Kirichenko, T. Radivilova, V. Bulakh, "Generalized approach to Hurst exponent estimating by time series. *Informatyka Automatyka Pomiary w Gospodarce i Ochronie Środowiska,* Volume 8, No. 1, pp.28-31, 2018. doi: 10.5604/01.3001.0010.8639

[13] J. W. Kantelhardt, "Fractal and Multifractal Time Series." *In: Meyers R. (eds) Mathematics of Complexity and Dynamical Systems*. Springer, New York, NY, 2012.

[14] R. G. Clegg, "A practical guide to measuring the hurst parameter." *International Journal of Simulation. Systems, Science & Technology*, Vol. 7, № 2. pp. 3–14, 2006.

[15] L. Kirichenko and T. Radivilova, "Analyzes of the distributed system load with multifractal input data flows," *In 2017 14th International Conference The Experience of Designing and Application of CAD Systems in Microelectronics (CADSM)*, Lviv, 2017, pp. 260-264. doi: 10.1109/CADSM.2017.7916130

[16] D.V. Ageyev, and M. T. Salah, "Parametric synthesis of overlay networks with self-similar traffic" *Telecommunications and Radio Engineering (English translation of Elektrosvyaz and Radiotekhnika)*, vol. 75, issue 14, pp. 1231-1241, 2016.

[17] Y. I. Daradkeh, L. Kirichenko and T. Radivilova, "Development of QoS Methods in the Information Networks with Fractal Traffic." *International Journal of Electronics and Telecommunications*, 2018, Vol. 64, No. 1, pp. 27-32. doi: 10.24425/118142

[18] D. Ageyev and D. Evlash, "Multiservice telecommunication systems design with network's incoming self-similarity flow," *2008 International Conference on "Modern Problems of Radio Engineering, Telecommunications and Computer Science" (TCSET)*, Lviv-Slavsko, 2008, pp. 403-405.

[19] L. Kirichenko, T. Radivilova and V. Bulakh, "Machine Learning in Classification Time Series with Fractal Properties." *Data,* vol.4(1) 5, pp.1-13, 2019. doi:10.3390/data4010005

[20] V. Bulakh, L. Kirichenko and T. Radivilova, "Classification of Multifractal Time Series by Decision Tree Methods," In *Proceedings of the 14th International Conference on ICT in Education, Research and Industrial Applications. Integration, Harmonization and Knowledge Transfer. Volume I: Main Conference*, vol.2105. Kyiv, Ukraine, 2018.

[21] V. Bulakh, L. Kirichenko and T. Radivilova, "Time Series Classification Based on Fractal Properties," *In Proceedings of the 2018 IEEE Second International Conference on Data Stream Mining & Processing (DSMP)*, Lviv, Ukraine, pp. 198–201, 2018. doi:10.1109/DSMP.2018.8478532.

[22] L. Kirichenko, T. Radivilova and A. S. Alghawli, "Mathematical simulation of self-similar network traffic with aimed parameters." *Annals Computer Science Sériés*, vol.11, tome 1, pp.17-22, 2013.

[23] L. Breiman, "Random Forests." *Machine Learning,* vol. 45(1), pp.5-32, 2001. doi: 10.1023/A:1010933404324

[24] S. Ioffe and C. Szegedy, "Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift," *In Proceedings of the 32nd International Conference on Machine Learning,* Lille, France, PMLR 37, pp.448-456, 2015. https://arxiv.org/abs/1502.03167

[25] D.P. Kingma and J. Ba, "Adam: A Method for Stochastic Optimization," *In: Proceedings of the 3rd International Conference on Learning Representations (ICLR)*, San Diego, USA (2015). https://arxiv.org/abs/1412.6980s