Bearing Faults Diagnosis and Classification Using Generalized Gaussian Distribution Multiscale Dispersion Entropy Features

Ragavesh Dhandapani
Department of Electrical and Communication Engineering
College of Engineering, National University of Science and Technology
PB 2322, CPO Seeb 111, Sultanate of Oman
ragavesh@nu.edu.om

Imene Mitiche
Department of Computing
Glasgow Caledonian University
70 Cowcaddens Road, Glasgow G40BA, UK
imene.mitiche@gcu.ac.uk

Scott McMeekin
School of Computing, Engineering and Built Environment
Glasgow Caledonian University
70 Cowcaddens Road, Glasgow G40BA, UK
scott.mcmeekin@gcu.ac.uk

Gordon Morison
Department of Computing
Glasgow Caledonian University
70 Cowcaddens Road, Glasgow G40BA, UK
gordon.morison@gcu.ac.uk

***Abstract*—Effective fault diagnosis of rolling bearings are vital for the reliable and smooth operation of industrial equipment. Early fault detection and diagnosis of rolling bearings are required to avoid catastrophic failures and financial losses. In this paper, we propose a new sophisticated Multiscale Dispersion Entropy (MDE) based feature that uses a nonlinear mapping approach using a Generalized Gaussian Distribution (GGD)–Cumulative Distribution Function (CDF). First of all, the proposed feature extraction method is used to extract the features from a raw 1-D vibration signal and the candidate feature of each vibration signal is selected by analysing the standard deviation of the features. Then, the features are used as input to a Multi-class Support Vector Machine (MCSVM) model for categorizing rolling bearing fault conditions. The findings demonstrate that the proposed method is better in terms of classification accuracy, precision, recall and F1–score as compared to other entropy feature driven classification models.**

***Index Terms*—Dispersion Entropy; Bearing Fault Classification; Generalized Gaussian Distribution; Multi-class Support Vector Machine; Multiscale Dispersion Entropy**

I. Introduction

Rotating machinery is extremely important in modern industry, and the failure of critical component such as rolling bearing in the machinery may lead to machine breakdown and catastrophic accidents. As a result, in recent decades, condition monitoring (CM) of rotating machinery has gained more popularity. Rolling bearings fault diagnosis is an important aspect in the CM plan in the industry. Analyzing the vibration signal is one of the most common methods for determining the condition of a rolling bearings. Recently, many types of Entropy measures are widely used in different domains to study the degree of randomness in the time series data, and extensively used in rolling bearings fault diagnosis system [1]–[11]. Some commonly used Entropy based measures in rolling bearings fault analysis are Sample Entropy (SE) proposed by [12] is used in [6]–[8], Permutation Entropy (PE) proposed by [13] is used in [1], [6], [9], Weighted PE proposed by [14] is used in [1] and Dispersion Entropy (DE) proposed by [15] is used in [6]. SE, PE, WPE and DE measures are used to quantify the signals on a single scale, which may be not conducive to the extraction of signal features.

Some types of multiscale entropy measures used in rolling bearings fault analysis are Multiscale Entropy (MSE) proposed by [16] used in [3], [8], [11], Multiscale Dispersion Entropy (MDE) and Refined Composite MDE (RCMDE) proposed by Azami et al. [17] were applied to extract features from the vibration signal in [2], [11] and [4], [5], [10], [11] respectively. MDE and RCMDE method addresses the shortcomings observed in MSE, such as undefined and unstable for short signals [17]. The Normal Cumulative Distribution Function (NCDF) maps the input signal from 0 to 1 in DE, MDE, and RCMDE methods. As suggested in [17] to use other nonlinear mapping approach, in this work, Generalized Gaussian Distribution (GGD) Cumulative Distribution Function (CDF) is selected instead of NCDF. In addition to the Gaussian distribution [20] parameters, GGD model has one more parameter, shape parameter β. The proposed novel Generalized Gaussian Distribution–Multiscale Dispersion Entropy (GGD–MDE) measure is used with different scale factors and shape factor to extract the multiscale features from the raw vibration signal to formulate a feature pool.

Feature selection is essential after feature extraction to reduce feature dimension and computational burden during the training process. As motivated by [18], the GGD–MDE feature cluster that has the lowest standard deviation (SD) among the different shape parameters is selected to form a primary feature pool. Finally, the primary feature pool is used to train a Multi-class Support Vector Machine (MCSVM) to

ISBN: 978-1-6654-6798-8

classify the bearing fault type with the fault severity level. The effectiveness of the proposed method is validated by using experimental data set. The main contributions of this paper are given as follows:

1) A novel GGD–MDE is proposed to extract the multi-scale entropy-based feature from 1-D signal.
2) A new rolling bearing fault diagnosis method based on GGD–MDE and MCSVM is formulated.

The structure of this paper is as follows: in section II, the proposed method and the theoretical background of GGD–MDE and MCSVM are briefly introduced. In section III, results and discussion is presented and in section IV, the conclusion is presented.

II. The Proposed Method

The simplified flowchart of the proposed rolling bearing fault diagnosis method is shown in Fig. 1 and the overall process of the proposed fault diagnosis approach is as follows:

A. Sensor Data Acquisition

Initially, acquire the vibration signal and label them suitably. In this study, the bearing vibration data set as shown in Fig. 2, available in [19] is used for training and testing purposes. Based on the condition of the bearing, the signals are labeled as 'Normal' for bearing without any fault, 'ORF007', 'ORF014' and 'ORF021' for bearing with 'Outer Race Fault' category, 'BF007', 'BF014', 'BF021' for bearing with 'Ball Fault' category and 'IRF007', 'IRF014', 'IRF021' for bearing with 'Inner Race Fault' category, totaling 10 (B) rolling bearing vibration signals. After indexing the data set, the feature of each data set is extracted as explained in the following section.

B. Feature Extraction

DE proposed by Rostaghi and Azami [15] with a new mapping method is resulted in Generalized Gaussian Distribution - Dispersion Entropy (GGD–DE) method. Further, GGD–DE is used in developing a novel GGD–MDE method. For feature extraction, initially, the vibration signals are divided into 'P' number of non-overlapped segments with 1K (1024) samples per segment. The feature-length 'P' is set as 110. Then, for each segment the GGD–MDE features are calculated for the different shape parameter β.

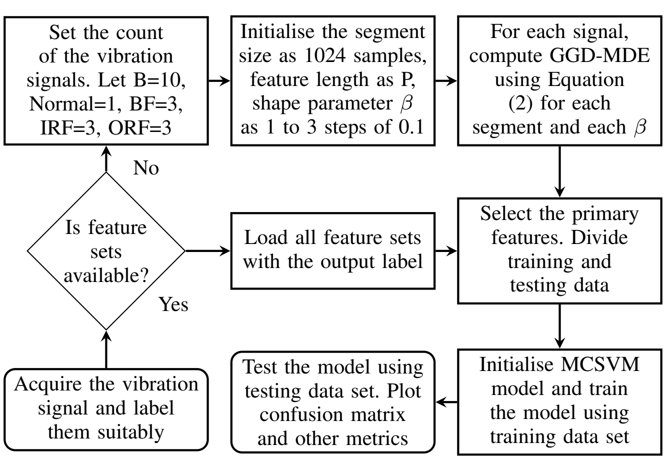

Fig. 1: Flowchart of the rolling bearing fault diagnosis method.

1) GGD–MDE: Assume that each segment of vibration signal is a univariate signal of length N: $\mathbf{u} = u_1, u_2, ..., u_N$, where $N = 1024$. The signal $\mathbf{u}$ is coarse-grained by dividing into $\left\lfloor \frac{N}{\tau} \right\rfloor$ non-overlapping segments with the length of scale factor τ, and then the average of each segment is calculated as proposed in [17]:

$$\mathbf{x}_j^{\tau} = \frac{1}{\tau} \sum_{b=(j-1)\tau+1}^{j\tau} u_b, \qquad 1 \le j \le \left\lfloor \frac{N}{\tau} \right\rfloor \tag{1}$$

where $\mathbf{x}_j^{\tau}$ denotes the j-th coarse-grained time series of $\mathbf{u}$. Finally, the Entropy value using GGD–DE is calculated for each coarse-grained signal by considering as $\mathbf{x}$ with the length *L*, the GGD–DE is calculated as follows:

Step 1: Initially, $\mathbf{x} = x_1, x_2, \cdots, x_L$ are mapped to $\mathbf{y} = y_1, y_2, \cdots, y_L$ from 0 to 1 using the CDF of GGD which is defined as:

$$\mathbf{y} = \frac{1}{2} + \operatorname{sgn}(\mathbf{x} - \mu) \frac{\hat{\gamma}\left[1/\beta, \left(\frac{|\mathbf{x}-\mu|}{\rho}\right)^{\beta}\right]}{2\Gamma(1/\beta)} \tag{2}$$

where $\hat{\gamma}$ denotes the lower incomplete gamma function, Γ denotes the gamma function, ρ is the scale parameter given

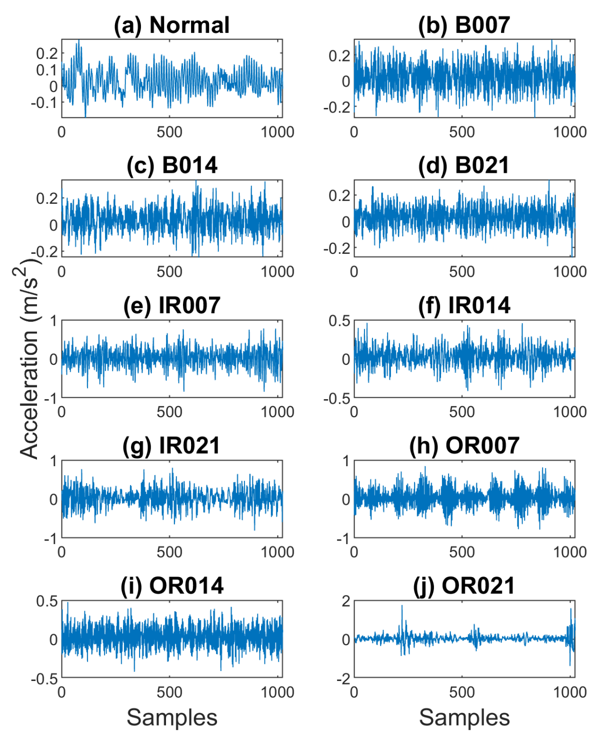

Fig. 2: Sample waveform of normal and fault conditions. 'Normal' for bearing without any fault, 'BF007', 'BF014', 'BF021' for bearing with 'Ball Fault' category, 'IRF007', 'IRF014', 'IRF021' for bearing with 'Inner Race Fault' category, 'ORF007', 'ORF014' and 'ORF021' for bearing with 'Outer Race Fault' category.

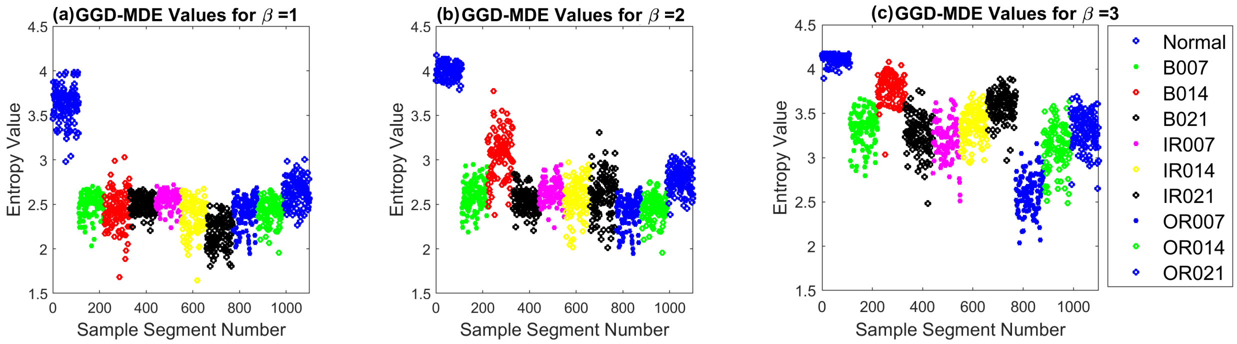

Fig. 3: GGD-RCMDE values of all working sample data set, when (**a**) β = 1 (**b**) β = 2 and (**c**) β = 3 at scale factor=15

by $\rho = \sigma\sqrt{\Gamma(1/\beta)/\Gamma(3/\beta)}$, μ is mean and σ is standard deviation of the signal $\mathbf{x}$, and $\mathbf{y}$ values are the probabilities that a random variable from the GGD, with μ and σ of the entire signal $\mathbf{x}$, is less than x value from the time series $\mathbf{x}$. The β determines shape of the distribution, e.g., the Laplacian for $\beta = 1$, the Gaussian for $\beta = 2$ and close to Uniform distribution by letting $\beta = \infty$ [20].

Step 2: Then, each y_j, where $j = 1, 2, \ldots, L$ is assigned a class from 1 to c by linear algorithms as follows:

$$z_j^c = \text{round}\left(c \cdot y_j + 0.5\right) \qquad j = 1, 2, \ldots, L \tag{3}$$

where z_j^c denotes the j-th member of the classified time series.

Step 3: The embedding vector $\mathbf{z}_i^{m,c}$ with dimension m and time delay d are generated using the following equation:

$$\mathbf{z}_i^{m,c} = \left\{z_i^c, z_{i+d}^c, \cdots, z_{i+(m-1)d}^c\right\} i = 1, 2, \cdots, L-(m-1)d \tag{4}$$

then, each time series $\mathbf{z}_i^{m,c}$ is mapped to a dispersion pattern $\pi_{v_0 v_1 \cdots v_{m-1}}$, where $v_0 = z_i^c$, $v_1 = z_{i+d}^c$, $\cdots$, $v_{m-1} = z_{i+(m-1)d}^c$. The total possible number of dispersion patterns is c^m.

Step 4: The relative frequency of each potential dispersion patterns $p\left(\pi_{v_0 v_1 \cdots v_{m-1}}\right)$ can be given by:

$$\frac{\#\ \left\{i | i \le L-(m-1)d, \mathbf{z}_i^{m,c} \text{ has type } \pi_{v_0 v_1 \ldots v_{m-1}}\right\}}{L-(m-1)d} \tag{5}$$

where $\#$ means the number of dispersion patterns of $\pi_{v_0 v_1 \cdots v_{m-1}}$ that is assigned to $\mathbf{z}_i^{m,c}$.

Step 5: Lastly, according to Shannon's definition of entropy, the GGD–DE value with embedding dimension *m*, the number of classes *c* and the time delay *d* is computed as follows:

$$\begin{aligned} &\text{GGD-DE}(\mathbf{x}, m, c, d, \beta) \\ &= -\textstyle\sum_{\pi=1}^{c^m} p\left(\pi_{v_0 v_1 \ldots v_{m-1}}\right) \cdot \ln\left(p\left(\pi_{v_0 v_1 \ldots v_{m-1}}\right)\right) \end{aligned} \tag{6}$$

For each signal, the GGD–MDE measure of each segment with the different shape parameter β $(1 \le 0.1 \le 3)$ are calculated and stored as a feature pool (**FP**) having the dimension given by $B \times P \times$ step count of $\beta \times \tau$.

The parameters used in this work are the embedding dimension *m* = 4, the number of classes *c* = 6, the time delay *d* = 1, scale factor τ = 15 as suggested in [15].

2) Feature Selection: Fig. 3 shows the GGD–MDE features for β = 1, 2 and 3 of 'P' segments for all 'B' vibration signals. For each signal, the GGD–MDE feature cluster from **FP** that has the lowest standard deviation (SD) among the different shape parameters is selected to form a primary feature pool. Then, the primary feature pool is divided for training and testing purposes based on the pre-defined ratio such as (40:60, 60:40, 80:20). Then, the training data is used to train the MCSVM classifier.

3) MCSVM Model: Support Vector Machine (SVM) is a binary classifier for learning and separating algorithm in pattern recognition tasks [21]. This study requires classification of different bearing faults, which is a multiclass problem. Directed Acyclic Graph (DAG), Binary Tree (BT), One-Against-One (OAO), and One-Against-All (OAA) are some of the MCSVM classification models [22]. In OAA, a specific category label in the input is trained as one class (positive class), while the rest of the input is trained as a negative class. As a result, each classifier predicts whether test data belongs to a specific class or not. Therefore, a *J* output class problem ($J \ge 2$), *J* number of binary SVM classifiers are constructed. In this work, OAA MCSVM with Gaussian kernel function is used as a classifier [21], [23].

III. Results and Discussions

To demonstrate the performance of the proposed method, we selected the experimental data set provided by Case Western Reserve University (CWRU) for training and testing phase of our work [19]. This data set has been widely used by many researchers, hence it is appropriate to benchmark our work against their findings.

A. Case Western Reserve University Bearing Data

The vibration data collected form the drive end bearings were used in our study. The test bearings such as normal and faulty bearings are mounted on the shaft of the motor. The artificial faults in the bearings were created at a particular location such as rolling element, inner raceway and outer raceway with fault diameters of 0.007, 0.014 and 0.021 inches. In this study, the drive end vibration data is collected at 12000 samples/second for 0 hp load is used. Few description of the data set is listed in Table I and more details can be found in [19].

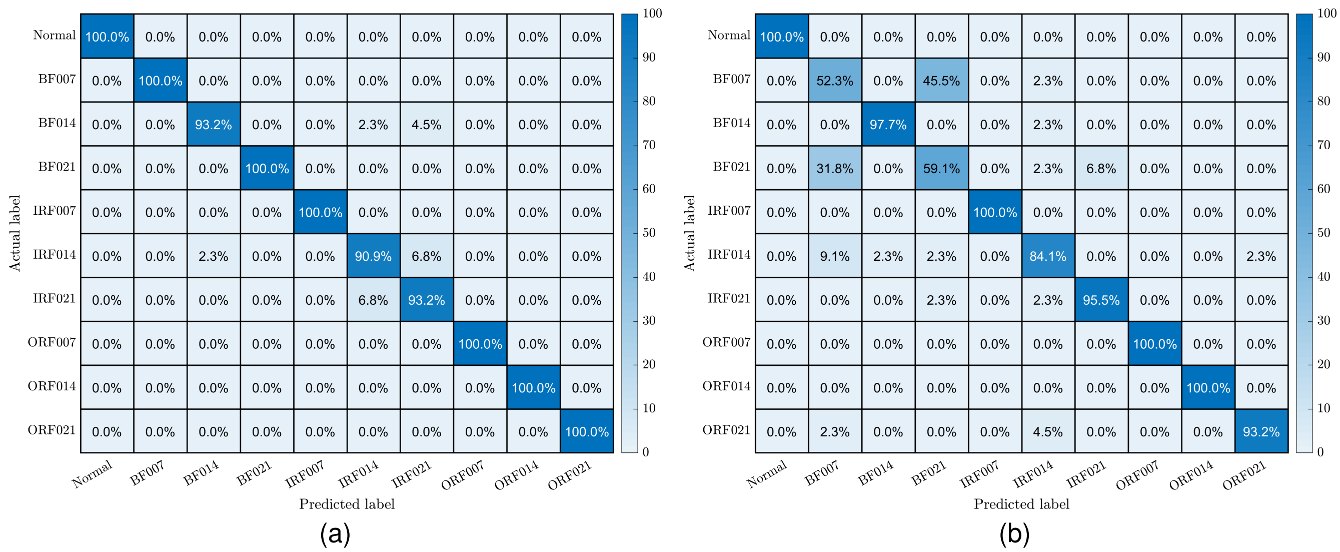

Fig. 4: The confusion matrix of (**a**) GGD-MDE method and (**b**) MDE (NCDF) method with 10 primary features and 0.6 training ratio.

B. Experimental Results

The Entropy values such as SamEn [6], [7], PerEn [6], Weighted PerEn [1], DE [4], [6] were discussed and presented in the literature for rolling bearings fault classification. As compared to the single scale DE measure [4], [6], multiscale entropy measures such as MSE [3], improved MDE [2] and RCMDE [4], [5] were shown better performance.

Entropy values of each segment for various multiscale methods such as MDE, MSE, RCMDE and GGD–MDE are extracted from the raw vibration signal. Entropy values of each method for all scales with the suitable class label forms a feature pool of that method. The following discussion shows how GGD–MDE features are used for rolling bearings fault diagnosis.

The feature pool comprises of 23100 $\times$ 15 data (10 signals $\times$ 110 segments $\times$ 21 steps of β $\times$ 15 scale factor). The SD of 110 segments for 21 shape parameters is computed for a signal, and the cohort with the lowest SD is chosen to construct a primary feature pool. This process is repeated for all signals, resulting in a primary feature pool of 1100 $\times$ 15 features. The β value at which the lowest SD of GGD–MDE values is attained to build a primary feature pool is shown in the distinct shape factor (DSF) column of Table I.

TABLE I: Description of Case Western Reserve University bearing data set [19].

Category	Fault size (inches)	Class label	File name	Standard deviation	DSF
Normal		Normal	97.mat	0.0462	2.7
	0.007	BF007	118.mat	0.1256	1.6
Ball Fault	0.014	BF014	185.mat	0.1491	3
	0.021	BF021	222.mat	0.0954	1.0
	0.007	IRF007	105.mat	0.0869	1.3
Inner Race	0.014	IRF014	169.mat	0.1763	3.0
Fault	0.021	IRF021	209.mat	0.1561	3.0
	0.007	ORF007	130.mat	0.1658	1.0
Outer Race	0.014	ORF014	197.mat	0.1304	1.0
Fault	0.021	ORF021	234.mat	0.1354	1.9

The primary feature pool data set is divided into two groups according to a pre-determined ratio, commonly referred as training and testing data set. This experiment considers three distinct training (0.4, 0.6, 0.8) and testing (0.6, 0.4, 0.2) ratios. Considering the total number of primary features, the training ratio 0.4 implies that 440 randomly selected features are used for training and the remaining 660 feature set are used for testing. The model is evaluated using the test data set after successful training, and performance metrics are produced using the confusion matrix. The Table II shows the performance metrics for all methods against the training ratio and the number of features. The following formulae is used to calculate the accuracy, precision, recall, and F1–score.

$$\begin{aligned} \text{Acc} &= \frac{TP+TN}{TP+FP+FN+TN} \\ \text{Pre} &= \frac{TP}{TP+FP} \\ \text{Rec} &= \frac{TP}{TP+FN} \\ \text{FS} &= 2 * \frac{\text{Precision.Recall}}{\text{Precision+Recall}} \end{aligned} \tag{7}$$

where TP is true positive (for example: actual label ‘IRF007’, predicted as ‘IRF007’), FP is false positive (actual label is not ‘IRF007’, predicted as ‘IRF007’), FN is false negative (actual label ‘IRF007’, predicted as other label) and TN is true negative which is the number of the accurate predictions of the other labels with respect to TP. The proposed GGD-MDE feature driven MCSVM classifier has the best classification accuracy (99.6%), better precision (0.98), recall (0.98) and F1-Score (0.98) as compared to the other methods proposed in [1], [2], [4] for the CWRU data set fault classification.

TABLE II: The average classification accuracy (Acc), precision (Pre), recall (Rec) and F1–score (FS) of MDE, RCMDE, MSE, and GGD-MDE techniques with various feature lengths and training ratios. The total number of features are 1100, training ratio of 0.6 implies that 660 are used for training and 440 are utilized for testing. The best classification metrics are in bold font.

Method	TR	Count of features=5				Count of features=10			
		Acc	*Pre*	*Rec*	*FS*	*Acc*	*Pre*	*Rec*	*FS*
MDE	0.4	96.9	0.85	0.83	0.82	97.3	0.87	0.86	0.86
	0.6	96.9	0.85	0.82	0.82	97.6	0.88	0.88	0.87
	0.8	97.1	0.85	0.84	0.83	97.6	0.88	0.88	0.87
RCMDE	0.4	97.7	0.89	0.88	0.87	98.7	0.94	0.94	0.94
	0.6	97.8	0.89	0.89	0.87	98.8	0.94	0.94	0.94
	0.8	98.0	0.90	0.90	0.88	98.9	0.95	0.95	0.95
MSE	0.4	96.0	0.80	0.79	0.79	97.2	0.86	0.86	0.86
	0.6	96.2	0.81	0.80	0.80	97.5	0.87	0.87	0.87
	0.8	96.4	0.82	0.81	0.81	97.4	0.87	0.87	0.87
GGD-MDE	0.4	99.3	**0.97**	**0.97**	**0.97**	**99.6**	**0.98**	**0.98**	**0.98**
	0.6	**99.4**	**0.97**	**0.97**	**0.97**	**99.6**	0.98	0.98	0.98
	0.8	99.3	**0.97**	**0.97**	**0.97**	**99.6**	0.98	0.98	0.98

IV. Conclusion

In this paper, a new GGD-MDE feature driven rolling bearings fault diagnosis method is proposed with the support of MCSVM. The proposed GGD mapping technique in MDE measure is a novel method that helps to extract multiple features from the raw vibration signal for a change of shape parameter β. A feature pool is formed by calculating the GGD–MDE value of the rolling bearing vibration data. Then, the primary feature pool is formed by analysing the SD of the features of a signal for each shape factor. Further, a part of the primary feature pool is used to train the MCSVM for classifying the normal and fault conditions and the remaining is used for testing. The proposed method with the new feature set is capable of classifying the bearing fault better than other multiscale entropy measures such as MSE, MDE and RCMDE. The statistical measures show that the proposed method is effective in bearing fault diagnosis as compared to the other methods considered in this work. Findings in this study are promising and this can be applied to other 1-D data sets such as ECG, EEG classification. The effect of segment length, embedding dimension m and number of classes c of the proposed GGD-MDE method can be evaluated.

Acknowledgment

The first author is extremely grateful to College of Engineering, National University for supporting this research work through their academic affiliate Glasgow Caledonian University.

References

[1] S. Zhou, S. Qian, W. Chang, Y. Xiao, and Y. Cheng, "A novel bearing multi-fault diagnosis approach based on weighted permutation entropy and an improved SVM ensemble classifier," *Sensors (Switzerland)*, vol. 18, no. 6, 2018.

[2] X. Yan and M. Jia, "Intelligent fault diagnosis of rotating machinery using improved multiscale dispersion entropy and mRMR feature selection," *Knowledge-Based Systems*, vol. 163, pp. 450–471, 2019.

[3] J. Zhang, Y. Zhao, X. Li, and M. Liu, "Bearing Fault Diagnosis with Kernel Sparse Representation Classification Based on Adaptive Local Iterative Filtering-Enhanced Multiscale Entropy Features," *Mathematical Problems in Engineering*, vol. 2019, 2019.

[4] W. Zhang and J. Zhou, "A comprehensive fault diagnosis method for rolling bearings based on refined composite multiscale dispersion entropy and fast ensemble empirical mode decomposition," *Entropy*, vol. 21, no. 7, 2019.

[5] J. Zheng and H. Pan, "Use of generalized refined composite multiscale fractional dispersion entropy to diagnose the faults of rolling bearing," *Nonlinear Dynamics*, vol. 101, no. 2, pp. 1417–1440, 2020.

[6] R. Li, C. Ran, B. Zhang, L. Han, and S. Feng, "Rolling bearings fault diagnosis based on improved complete ensemble empirical mode decomposition with adaptive noise, nonlinear entropy, and ensemble SVM," *Applied Sciences (Switzerland)*, vol. 10, no. 16, 2020.

[7] W. Jiao, G. Li, Y. Jiang, R. Baim, C. Tang, T. Yan, X. DIng, and Y. Yan, "Multi-Scale Sample Entropy-Based Energy Moment Features Applied to Fault Classification," *IEEE Access*, vol. 9, pp. 8444–8454, 2021.

[8] K. Zhu, X. Song, and D. Xue, "A roller bearing fault diagnosis method based on hierarchical entropy and support vector machine with particle swarm optimization algorithm," *Measurement: Journal of the International Measurement Confederation*, vol. 47, no. 1, pp. 669–675, 2014.

[9] V. Vakharia, V. K. Gupta, and P. K. Kankar, "A multiscale permutation entropy based approach to select wavelet for fault diagnosis of ball bearings," *JVC/Journal of Vibration and Control*, vol. 21, no. 16, pp. 3123–3131, 2015.

[10] Q. Wang, Y. Xiao, S. Wang, W. Liu, and X. Liu, "A Method for Constructing Automatic Rolling Bearing Fault Identification Model Based on Refined Composite Multi-Scale Dispersion Entropy," *IEEE Access*, vol. 9, pp. 86 412–86 428, 2021.

[11] J. Lv, W. Sun, H. Wang, and F. Zhang, "Coordinated approach fusing rcmde and sparrow search algorithm-based svm for fault diagnosis of rolling bearings," *Sensors*, vol. 21, no. 16, 2021.

[12] J. S. Richman and J. R. Moorman, "Physiological time-series analysis using approximate entropy and sample entropy," *American Journal of Physiology-Heart and Circulatory Physiology*, vol. 278, no. 6, pp. H2039–H2049, jun 2000.

[13] C. Bandt and B. Pompe, "Permutation Entropy: A Natural Complexity Measure for Time Series," *Physical Review Letters*, vol. 88, no. 17, p. 4, 2002.

[14] B. Fadlallah, B. Chen, A. Keil, and J. Príncipe, "Weighted-permutation entropy: A complexity measure for time series incorporating amplitude information," *Physical Review E*, vol. 87, no. 2, p. 022911, feb 2013.

[15] M. Rostaghi and H. Azami, "Dispersion Entropy: A Measure for Time-Series Analysis," *IEEE Signal Processing Letters*, vol. 23, no. 5, pp. 610–614, 2016.

[16] M. Costa, A. L. Goldberger, and C. K. Peng, "Multiscale entropy analysis of biological signals," *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics*, vol. 71, no. 2, pp. 1–18, 2005.

[17] H. Azami, M. Rostaghi, D. Abasolo, and J. Escudero, "Refined Composite Multiscale Dispersion Entropy and its Application to Biomedical Signals," *IEEE Transactions on Biomedical Engineering*, vol. 64, no. 12, pp. 2872–2879, 2017.

[18] A. Yousefpour, H. N. A. Hamed, U. H. H. Zaki, and K. A. M. Khaidzir, "Feature subset selection using mutual standard deviation in sentiment mining," *2017 IEEE Conference on Big Data and Analytics, ICBDA 2017*, vol. 2018-Janua, pp. 13–18, 2018.

[19] "Case Western Reserve University Bearing Data Center Website. Available online: http://csegroups.case.edu/ bearingdatacenter/home (accessed on 16 January 2021)." [Online]. Available: http://csegroups.case.edu/ bearingdatacenter/home

[20] S. Yu, A. Zhang, and H. Li, "A Review of Estimating the Shape Parameter of Generalized Gaussian Distribution," vol. 21, pp. 9055–9064, 2012.

[21] V. Vapnik N., "The Nature of Statistical Learning," 1995. [Online]. Available: https://ci.nii.ac.jp/naid/10020951890

[22] A. Widodo and B. S. Yang, "Support vector machine in machine condition monitoring and fault diagnosis," *Mechanical Systems and Signal Processing*, vol. 21, no. 6, pp. 2560–2574, 2007.

[23] C. Cortes and V. Vapnik, "Support-vector networks," *Machine Learning*, vol. 20, no. 3, pp. 273–297, sep 1995.