

Article

A Machine Learning Approach for Gearbox System Fault Diagnosis

Jan Vrba ^{1,*} , Matous Cejnek ^{2,*}  and Jakub Steinbach ¹  and Zuzana Krbcova ¹ 

¹ Department of Computing and Control Engineering, Faculty of Chemical Engineering, University of Chemistry and Technology, Technicka 5, 166 28 Prague, Czech Republic; Jakub.Steinbach@vscht.cz (J.S.); Zuzana.Krbcova@vscht.cz (Z.K.)

² Department of Instrumentation and Control Engineering, Faculty of Mechanical Engineering, Center of Advanced Aerospace Technology, Czech Technical University in Prague, Technicka Street 4, 166 07 Prague, Czech Republic

* Correspondence: jan.vrba@vscht.cz (J.V.); matous.cejnek@fs.cvut.cz (M.C.)

Abstract: This study proposes a fully automated gearbox fault diagnosis approach that does not require knowledge about the specific gearbox construction and its load. The proposed approach is based on evaluating an adaptive filter's prediction error. The obtained prediction error's standard deviation is further processed with a support-vector machine to classify the gearbox's condition. The proposed method was cross-validated on a public dataset, segmented into 1760 test samples, against two other reference methods. The accuracy achieved by the proposed method was better than the accuracies of the reference methods. The accuracy of the proposed method was on average 9% higher compared to both reference methods for different support vector settings.

Keywords: fault diagnosis; gearbox; adaptive filter; NLMS algorithm; support vector machine



Citation: Vrba, J.; Cejnek, M.; Steinbach, J.; Krbcova, Z. A Machine Learning Approach for Gearbox System Fault Diagnosis. *Entropy* **2021**, *23*, 1130. <https://doi.org/10.3390/e23091130>

Academic Editors: Claude Delpha and Demba Diallo

Received: 5 August 2021

Accepted: 27 August 2021

Published: 30 August 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The gearbox fault detection problem has become an intensively studied topic in the last few decades. Early detection of possible gearbox faults (or rotating machinery in general) can increase the operational safety of a device, as it can reduce the costs of maintenance and prevent total failure. To detect a possible fault of a gearbox, the vibrations of the gearbox are typically measured, often by using multiple sensors at once. The damaged teeth in the gearbox produce force impulses that are usually reflected in the vibration signal. However, evaluating of the vibration signal is difficult. We can view the underlying process of generating the vibration signal as strongly non-linear, non-stationary, and non-Gaussian. The faults produce a signal with energy distributed over various frequencies, which makes successful detection even more difficult. Another issue that arises with the analysis of gearbox vibration signals is that every single gearbox produces a unique signal, so the approach and settings that are successful with some gearboxes may completely fail for different ones. That is probably the reason many other published methods were not validated and evaluated with appropriate datasets. Due to the popularity of this topic, many methods utilizing different approaches have been developed.

As was already mentioned, the signal of a faulty gearbox or rotating machinery has a different spectrum than the faultless one. The fast Fourier transform [1] (FFT) is often used, because it is the most natural tool for studying the frequency spectrum. In the study [2], the authors utilized the DSP-based FFT analyzer that takes advantage of pattern matching techniques to detect faults in a rotating machine. The study [3] evaluated the usability of FFT for predictive maintenance of electrical rotating machines in connection with ISO 2372. In [4] the authors used sparse filtering to extract frequency domain features, classified by the softmax regression classifier, and obtained the output diagnosis results. In [5], the authors suggest using the multiscale chirplet path pursuit algorithm to approximate the best order

of the fractional Fourier transform (FrFT) by estimating the instantaneous frequency of the signal component with the largest energy. Then the FrFT spectrum of this component is analyzed and the fault is detected by the frequency sideband evaluation. In [6], the authors combined corrected multiresolution FT with discrete wavelet transform to investigate the vibrations of a gearbox and current transients of a connected DC generator. A comparison of FT and continuous wavelet transform for gearbox fault diagnosis is presented in [7].

There are approaches that are focused on residual signal analysis, such as the utilization of the auto-regressive model [8] or an auto-regressive model with exogenous input [9], where the residual signal is processed and the fault is detected by its features. The study [10] used a neural network to obtain the residual signal which, after Hilbert transform, provided significant information about the gearbox faults.

In some studies, only empirical mode decomposition (EMD) [11] without a Hilbert–Huang transform was used [12,13]. The combination of the Hilbert transform and EMD was presented in the study [14]. The EMD seems an interesting approach to signal evaluation due to its time complexity, as both EMD and FFT have time complexity $\mathcal{O}(n \log n)$ [15].

Various methods based on adaptive filtering have been developed over decades. In [16], the authors compared least mean squares (LMS) with linear prediction, spectral kurtosis, and fast block LMS to detect the bearing defect in a gearbox via spectral analysis (note that one of the first uses of LMS in condition monitoring was presented in [17]). They extended their work by comparisons with self-adaptive noise cancellation in [18], and claimed that LMS can, as mentioned in the previous study [16], detect a fault earliest. Another adaptive approach, namely, the adaptive line enhancer, was used in [19]. There were also multiple publications dedicated to the adaptive Schur filter (ASF). The ASF consists of several sections which are described by time-dependent reflection coefficients. Based on the forward prediction error and backward prediction error, the reflection coefficient is calculated for each section [20]. In publication [21], the authors proposed a framework for fault detection based on the changes of the prediction error of the Schur filter. The approach based on monitoring of changes of Schur filter coefficients was presented in [22]. This approach was extended in [23]. An approach to detecting fatigue tooth cracks in a wind turbine gearbox based on the adaptive Morlet wavelet filter was presented in [24]. The study [25] presented a combination of an adaptive noise reducer-based Gaussian reference signal technique with the a one-against-one multi-class support vector machine to detect various fault types in a gearbox. The self-adaptive noise cancellation method with nonlinear adaptive filter using a kernel least mean squares algorithm was presented in [26]. Another approach that is based on the adaptive regression splines method and trend change detection was presented in [27]. In [28], the authors proposed a new impulse energy indicator. They utilized an adaptive filter for signal separation, wavelet packet decomposition, and the combination of RMS and kurtosis to select the optimum filter band which indicates the fault in the bearing of the gearbox.

A unique approach based on the estimation of the cointegration factor of a vibration signal to detect the fault of a gearbox was presented in article [29]. Quantitative vibration analysis of bearing faults is exhaustively presented in [30], where the authors present a dynamic model of rolling element bearings and provide simulation results for a specific fault.

Recently, multiple methods using deep learning have emerged. In [31] the authors proposed to use an augmented deep sparse autoencoder to process the raw vibration signal. Study [32] avoided the need for a large dataset to teach a deep learning model by using a stacked sparse autoencoder that processes time-frequency images. Another approach that was presented in [33] uses multimodal deep support vector classification in combination with a Gaussian–Bernoulli deep Boltzmann machine. In [34], the authors combined improved particle swarm optimization, variational mode decomposition, and an improved convolutional neural network to process a signal spectrum and composite fault signal. In [35] the authors processed vibration, acoustic, and torque signals via discrete wavelet transform to obtain initial features for deep neural networks. The usage of convo-

lutional neural networks was also presented in [36]. Acoustic-based diagnosis based on a multiscale convolutional learning structure and an attention mechanism was presented in [37]. An interesting approach based on image processing was introduced in [38], where images with signal frequency spectra obtained via variational mode decomposition are used as inputs for a convolutional neural network. The study [39] presented the use of a deep random forest fusion technique to fuse acoustic emission and vibratory signals to detect various gearbox faults. In [40], the authors compared long-short-term memory and bi-directional long-short-term memory (LSTM) models for gearbox health monitoring. In [41], the authors transformed the vibration signal into an image-like simplified health data map that visualized a tooth-wise fault of the gearbox. This image was then processed by a convolutional neural network, and the remaining domain shift problem was solved via maximum classifier discrepancy. A diagnostic method based on time-frequency representation and deep reinforcement learning was presented in [42]. A diagnostic method based on bidirectional convolutional LSTM networks is presented in [43]. The authors claimed that their architecture can solve the problem of extracting spatial and temporal features simultaneously without losing any information. The study [44] introduced 1D deep convolutional transfer learning to process a torque measurement and estimate the health state of the gearbox. A deep morphological convolutional neural network for vibration signal processing was introduced in the article [45]. A new special type of CNN—the multiscale fusion global sparse network—for gearbox fault diagnosis, was proposed in [46]. Another unique neural network architecture, AKRnet, utilizing attentive kernel residual learning for feature learning of gearbox vibration signals, was presented in [47]. In [48], the authors proposed a fault diagnosis system that combines ResNet [49] with wavelet transform, and showed that their hybrid attention-based method improves ResNet's performance. A novel deep neural network which combines EMD, LSTM, and particle swarm optimization was presented in the study [50]. A method based on the usage of a two-class nonnegative matrix factorization network was proposed in [51]. There are many more applications of deep learning techniques in gearbox fault diagnosis that were published recently, which we are aware of but not mentioning here due to the scope of this article. In the study [52], the authors applied self-organizing maps with kurtosis criterion obtained via variational mode decomposition.

More information about gearbox fault detection approaches can be found in the following review papers: [53–55].

Most of the studies mentioned above were mainly qualitative, and some of their methods require expert opinions to conclude on the gearbox's condition. However, in our study, we focused on a simple, robust, and fully automated solution of gearbox fault detection without prior knowledge of the operation or measurement details, and without a need for a large training dataset. Those factors make it different to many of the other deep learning-based methods. The proposed method features a multiscale approach and can utilize a custom number of parallel sensors attached to the gearbox. The evaluation criteria were assessed using measurement data, and the proposed approach was cross-validated.

2. Materials and Methods

Our proposed method is based on utilizing the error prediction of an adaptive filter. The description of the adaptive filter is presented in Section 2.2. The adaptive filter coefficients are adapted via NLMS algorithm (see Section 2.3). Detailed description of the proposed descriptor is in Section 2.4. This descriptor is evaluated via support vector machine (SVM) and 10-fold cross-validation is performed (see Sections 2.6 and 2.7). Information about the used dataset is provided in Section 2.1 and overall experiment description in Section 2.8. Results of the proposed approach are compared with multiple methods in Section 3, those reference methods are introduced in Section 2.5.

2.1. Dataset

A publicly available dataset already presented in the paper [56] was used in this work. The dataset includes vibration measurements from healthy and broken gearboxes under various loads and a constant rotating speed at 30 Hz. The measurements were recorded via SpectraQuest's Gearbox Fault Diagnostics Simulator.

The original whole data \mathbf{x}_s contain a time series of various lengths. This unbalance was corrected with data segmentation. The original dataset was segmented into $K = 1000$ data point-long segments. For every condition (broken/healthy) and every load (0–90% with step 10%), $P = 88$ segments were formed. Leftover data from the longer time series were not used. The resulting balanced dataset consists of 1760 segments. Every segment contains $J = 4$ time series measured by sensors placed on different places (4 sensors). Therefore, the data segment can be described as

$$\mathbf{x}_j^p = [x_{sj}(p \cdot K - K), x_{sj}(p \cdot K - K + 1), \dots, x_{sj}(p \cdot K - 1)] \quad (1)$$

Note that index $j \in \{1, 2, 3, 4\}$ represents the sensor ID and $p \in \{1, 2, \dots, N\}$ is the segment number. Note the total number of segments available $N = 1760$ is:

$$N = P \cdot n_{condition} \cdot n_{loading} \quad (2)$$

where $n_{condition} = 2$ (broken, healthy) and $n_{loading} = 10$ (loading from 0 to 90% with 10% step).

According to our results, every sensor provides a different amount of information about the health status of the gearbox. Furthermore, the information from the independent sensors is not fully correlated, and thus the sensors can be used in a complementary fashion.

2.2. Adaptive Filter

The scheme that depicts the adaptive filtration problem is in Figure 1, where $x(k) \in \mathbb{R}$ is the input signal, $v(k) \in \mathbb{R}$ represents additive noise, $y(k) \in \mathbb{R}$ is measured output of the system, $\hat{y}(k) \in \mathbb{R}$ is the output of the adaptive filter and $e(k) \in \mathbb{R}$ represents the error of the filter (prediction error). In our proposed method, we have used the linear adaptive filter with finite impulse response. The output of the filter at discrete time index $k \in \mathbb{Z}$ is given by

$$\hat{y}(k) = \sum_{i=0}^N w_i(k)x(k-i) \quad (3)$$

where $w_i(k) \in \mathbb{R}$ is the value of the i -th adaptive weight, $N + 1$ is the number of adaptive weights and $x(k - i)$ represents the delayed sample. The output of the FIR filter (Equation (3)) in vector notation can be written as

$$\hat{y}(k) = \mathbf{x}(k)^T \cdot \mathbf{w}(k) \quad (4)$$

where $\mathbf{w}(k)$ is the vector of adaptive weights given as

$$\mathbf{w}(k) = [w_0(k), \dots, w_N(k)] \quad (5)$$

and $\mathbf{x}(k)$ represents the input vector given as

$$\mathbf{x}(k) = [x(k), \dots, x(k - N)]. \quad (6)$$

Note that the adaptive weights were updated via NLMS algorithm (Section 2.3) with every sample obtained.

The block schema of the filter is depicted in Figure 2. The block with z^{-1} represents unit time delay (if we consider Z-transform notation).

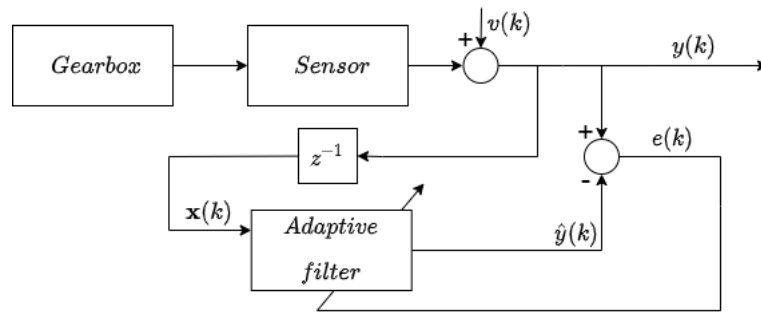


Figure 1. Block schema of adaptive filtration process.

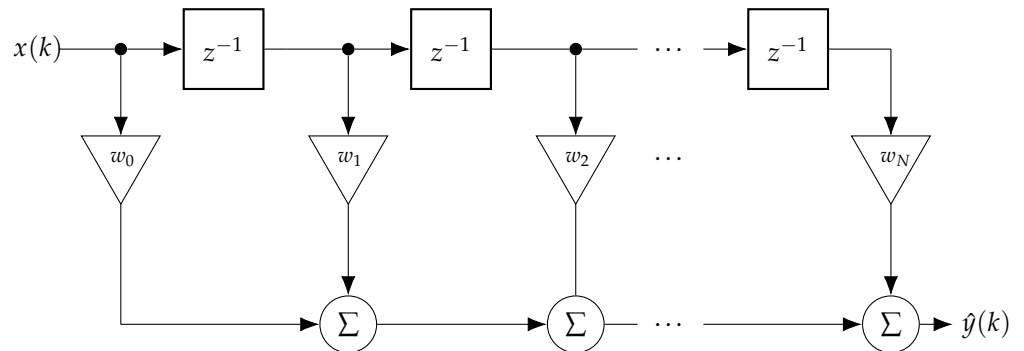


Figure 2. Block schema representation of a digital adaptive filter.

2.3. The NLMS Algorithm

The proposed method is based on the error evaluation of a normalized-last-mean-square (NLMS) adaptive filter [57] in predictive settings.

The NLMS algorithm is a modification of the LMS (stochastic gradient descent). The LMS weight adaptation is given as follows.

$$\mathbf{w}(k + 1) = \mathbf{w}(k) + \Delta \mathbf{w}(k), \tag{7}$$

where $\Delta \mathbf{w}(k)$ is

$$\Delta \mathbf{w}(k) = \frac{1}{2} \mu \frac{\partial e^2(k)}{\partial \mathbf{w}(k)} = \mu \cdot e(k) \cdot \mathbf{x}(k), \tag{8}$$

where $\mu \in \mathbb{R}$ is the learning rate (step size) and $e \in \mathbb{R}$ is error, which is defined as

$$e(k) = y(k) - \hat{y}(k). \tag{9}$$

The NLMS adaptation rule is given as follows:

$$\Delta \mathbf{w}(k) = \eta(k) \cdot \mathbf{w}(k) \cdot e(k), \tag{10}$$

where $\eta(k)$ is an actual learning rate normalized with $\|\mathbf{x}(k)\|^2$ (input signal power) as follows:

$$\eta(k) = \frac{\mu}{\epsilon + \|\mathbf{x}(k)\|^2}. \tag{11}$$

where $\epsilon \in \mathbb{R}$ is a small positive constant (regularisation term) introduced to preserve stability for inputs close to zero. The NLMS with ϵ is also called ϵ -NLMS. The NLMS algorithm is stable if

$$0 \leq \mu \leq 2 + \frac{2\epsilon}{\|\mathbf{x}(k)\|^2}, \tag{12}$$

or for the case without a regularization term: ϵ

$$\mu \in \langle 0, 2 \rangle. \quad (13)$$

The optimal learning rate is affected by the properties of additive noise $v(k)$. In the case that the additive noise is uncorrelated with input signal \mathbf{x} , the optimal learning rate is given as

$$\mu_{optimal} = \frac{E[|\hat{y}(k) - \tilde{y}(k)|^2]}{E[|e(k)|^2]} \quad (14)$$

The adaptive filter prediction error is obtained for every predicted data point in step-by-step predictive settings. This operation is done for every sensor independently. In this case, an input of the filter $\mathbf{x}(k)$ consists of previously measured values \mathbf{y} as follows:

$$\mathbf{x}(k) = [y(k-1), \dots, y(k-n)]. \quad (15)$$

where the filter length n directly represents the number of historical values used for prediction. In this study, the filter size was set to $n = 10$, because this setting yielded the best results.

2.4. Proposed Descriptor

To detect a gearbox fault, we propose a descriptor based on the prediction error of the adaptive filter adapted via NLMS algorithm. This data segment \mathbf{x}_j^p is processed by an adaptive filter that utilizes NLMS algorithm to obtain the filter prediction error \mathbf{e}_j^p . Assume the corresponding segment of filter prediction error as

$$\mathbf{e}_j^p = [e_j^p(0), e_j^p(1), \dots, e_j^p(K)] \quad (16)$$

The resulting descriptor for the whole segment of filter error \mathbf{e}_p^j (9) is evaluated as follows.

$$c_{error} = \sqrt{E[(\mathbf{e}_j^p)^2] - (E[\mathbf{e}_j^p])^2}, \quad (17)$$

where E stands for the expected value. This descriptor represents the actual effort of the adaptive filter made to follow the target signal. This effort can be understood as a novelty in the data, or irregularity in the measured signal. Relation between irregularity in the data and the health condition of the gearbox is the main idea behind the proposed method. Note that in our study we processed only nonoverlapping segments.

2.5. Reference Methods

To present a suitable challenge for the proposed method, two other reference methods were used to evaluate the health status of the gearbox.

2.5.1. Standard Deviation of the Raw Data

The first reference method is based on evaluation of the raw measured data standard deviation as the direct feature for the gearbox health assessment:

$$c_{plain} = \sqrt{E[\mathbf{y}^2] - (E[\mathbf{y}])^2}. \quad (18)$$

This method represents the most intuitive and computationally cheap way to measure irregularities in data. However, this approach is usable only if the standard deviation represents the underlying distribution.

2.5.2. Standard Deviation of the First IMF

Various frequency analysis methods are common tools for diagnosing faults in rotating machinery. Therefore, the second reference method used is based on frequency analysis of data. Due to the nonstationary nature of the gearbox vibration measurements, the Hilbert–Huang transform (HHT) [58] is used more often than the Fourier transform (FT) [1]. The fundamental part of the HHT is the empirical mode decomposition (EMD) method [11]. The EMD breaks down signals into components. In some studies, the EMD without the HHT is used [12,13]. Using the EMD method, any complicated dataset can be decomposed into a finite and often small number of components—intrinsic mode functions (IMFs). Any IMF represents a generally simple oscillatory mode that can be understood as a counterpart to the simple harmonic function. An IMF is any function with the same number of extrema and zero crossings, whose envelopes are symmetric with respect to zero. An example of a gearbox signal's sample IMFs is presented in Figure 3. The process of an IMF extraction is called sifting. The sifting process can be described as:

1. Create upper and lower envelope—connect local minima/maxima by a cubic spline line.
2. Get the first component (\mathbf{h}_1) from original signal and mean value (\mathbf{m}_1) of lower and upper envelopes:

$$\mathbf{x} - \mathbf{m}_1 = \mathbf{h}_1. \quad (19)$$

The envelopes and local extremes of a gearbox signal sample are shown in Figure 4.

3. Repeat the previous step i times:

$$\mathbf{h}_{1(i-1)} - \mathbf{h}_{1i} = \mathbf{h}_{1i}. \quad (20)$$

4. Obtain the first IMF

$$\mathbf{imf}_1 = \mathbf{h}_{1(i)} \quad (21)$$

5. Repeat previous steps to get other IMFs.

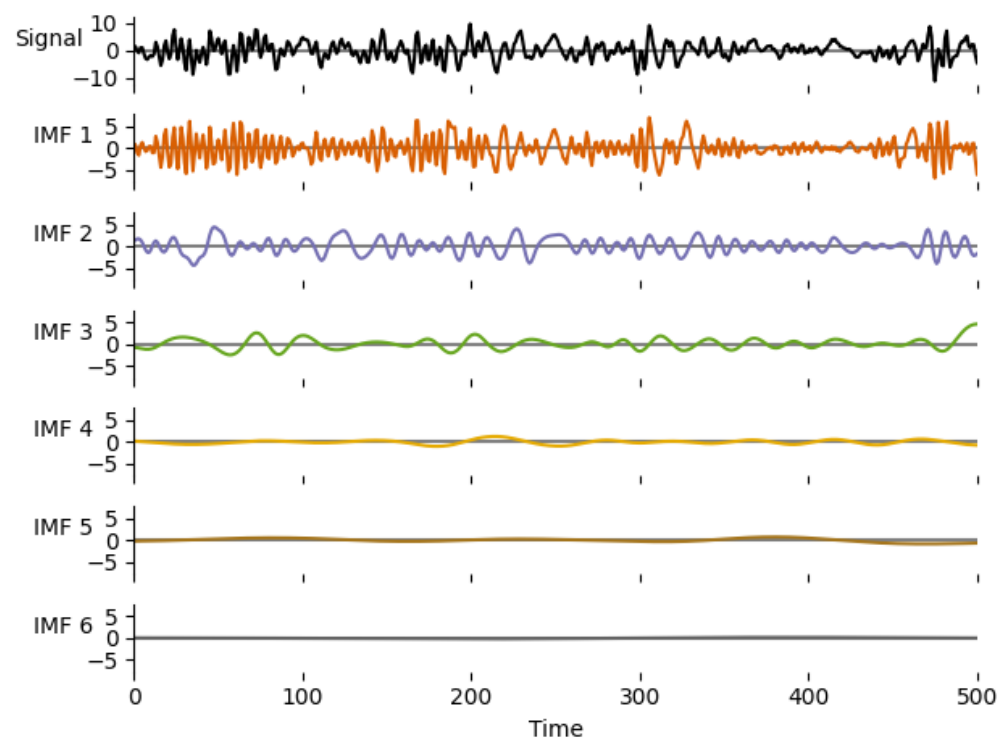


Figure 3. Resulting IMF of a gearbox signal sample. Note: an IMF is not a harmonic wave and the higher IMFs contain waves with generally lower frequency—the last IMF represents the slowest trend in the signal.

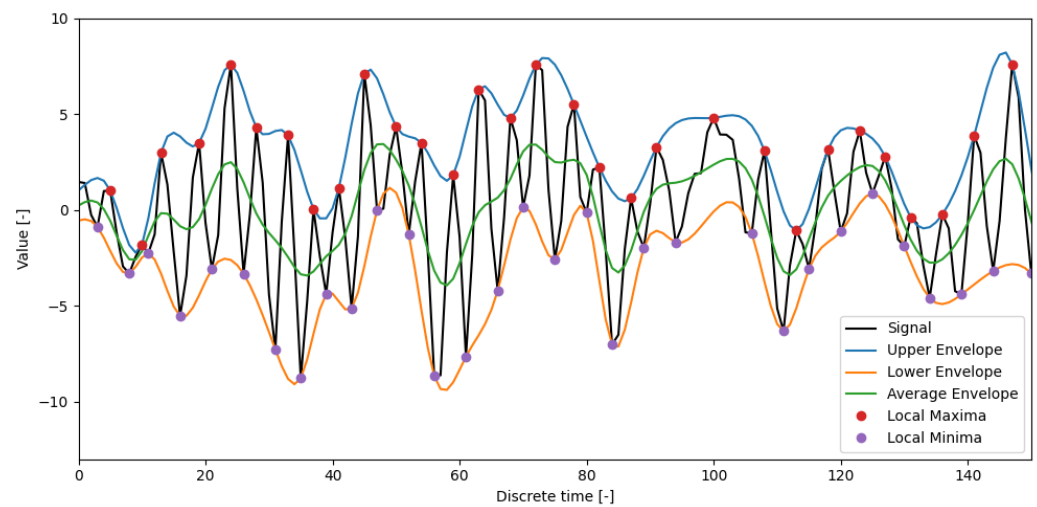


Figure 4. One step of the sifting process applied to a gearbox signal sample. Upper and lower envelopes are cubic splines connecting the local extremes.

The number of sifting steps (j) is determined via the stoppage criterion. There are four known stoppage criteria for the sifting process: the standard deviation criterion; the number criterion; the threshold method; energy difference tracking. The implementation of EMD used in this study utilizes the standard deviation criterion. The sifting process stops when the standard deviation is smaller than a pre-supplied value. The standard deviation of the sifting step is determined as [40]:

$$SD_i = \sum_{k=0}^K \frac{|\mathbf{h}_{i-1}(k) - \mathbf{h}_i(k)|^2}{\mathbf{h}_{i-1}^2(k)}. \quad (22)$$

In this work, the EMD was used to filter out the slower frequencies from the data to enhance the standard deviation change present in the original data. Then, the standard deviation was evaluated utilizing the first intrinsic mode function (IMF):

$$c_{emd} = \sqrt{E[\mathbf{imf1}^2] - (E[\mathbf{imf1}])^2}. \quad (23)$$

The first IMF was chosen because it had the greatest influence on the final accuracy in this given experiment setup. This frequency-based data selection enhances the irregularity related to the gearbox condition and thus improves any evaluation.

2.6. Decision Making via SVM

All studied methods (proposed and reference ones) extract a single scalar (standard deviation) as the descriptor of the whole sensor measurement. However, the measured data contain four different sensor measurements. Every sensor contains slightly different information. To make a prediction on gearbox health state from those 1–4 available descriptors, the SVM [59] is used. The used SVM dataset of k points was formed as

$$(\mathbf{d}_1, l_1), \dots, (\mathbf{d}_k, l_k), \quad (24)$$

where every \mathbf{d}_i is a vector of 1–4 data segment descriptors and l_i is the data segment class (healthy/faulty). Assume the hyper-plane that divides the group of points into two classes, takes the form

$$\mathbf{a}^T \mathbf{d} + b = 0, \quad (25)$$

Assume that the support vector equations representing the healthy/faulty class are given as

$$-\mathbf{a}^T \mathbf{d} + b \geq \alpha \text{ for } l_i == +1 \quad (26)$$

$$-\mathbf{a}^T \mathbf{d} + b \leq \alpha \text{ for } l_i = -1 \quad (27)$$

where $+1$ and -1 represent the healthy class and faulty class. To find the optimal hyper-plane, we have to solve the following minimization problem:

$$\min_{a,b} \frac{1}{2} \|a\|^2$$

subject to $l_i(a \cdot x + b) - 1 \geq 0$, for $i = 1, \dots, k$ (28)

Note that $\frac{b}{\|a\|}$ determines the offset of the hyper-plane from the origin along the normal vector \mathbf{a} . A linear kernel is used in the SVM algorithm implementation [60].

2.7. N-Fold Cross-Validation

Exhaustive 10-fold cross-validation was used to obtain the accuracies of the proposed and reference methods. Before splitting the data into ten same size chunks, the segments were shuffled to balance the healthy vs. broken segment ratio in every chunk. In every iteration of cross-validation, one chunk was used for testing and nine chunks were used for training. The average accuracy was determined over all testing chunks. The scheme of the cross-validation process is depicted in Figure 5.

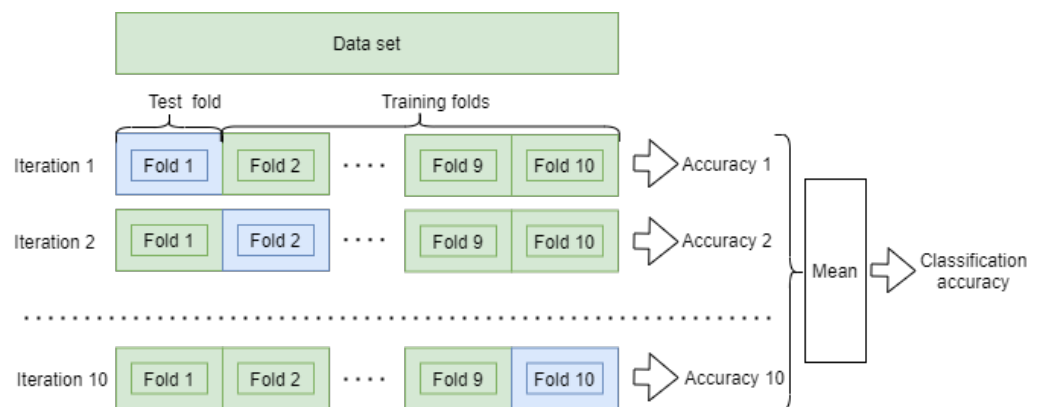


Figure 5. Schematics of the 10-fold cross-validation process used in this study.

2.8. Experimental Framework Overview

The core of the proposed method is the approach of extracting features from the raw data. The extracted features can be evaluated with any classifier. In this study, the well-known support vector machine (SVM) [59] was used. The classification outputs (gearbox condition) of the proposed method and the reference methods were validated via n-fold cross-validation. An overview of the experiment is displayed in Figure 6.

In this study, we compare the proposed method with two reference methods. The data shapes and organization of the methods in the testing framework are displayed in Figure 7.

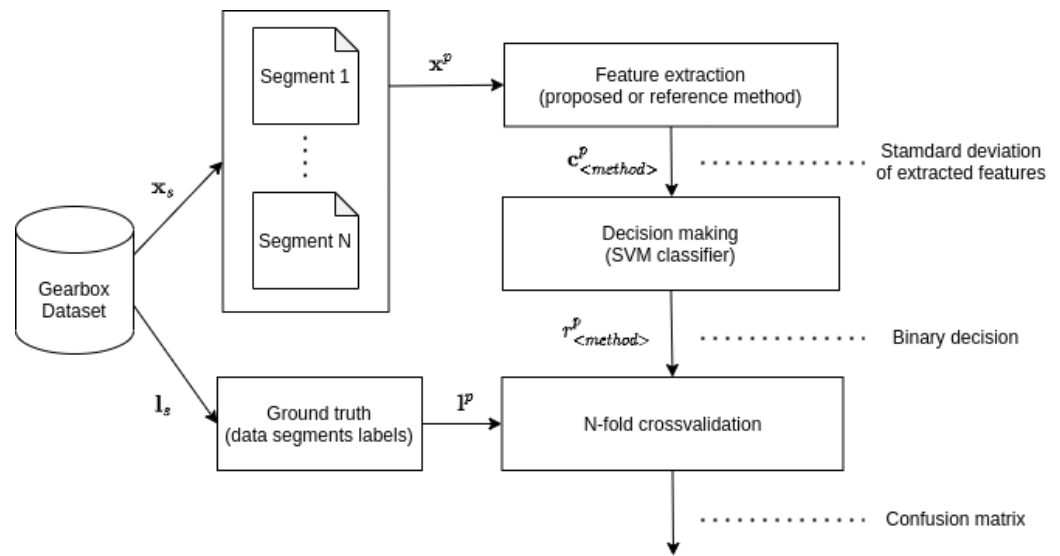


Figure 6. Experimental design—from raw data to final accuracy measure.

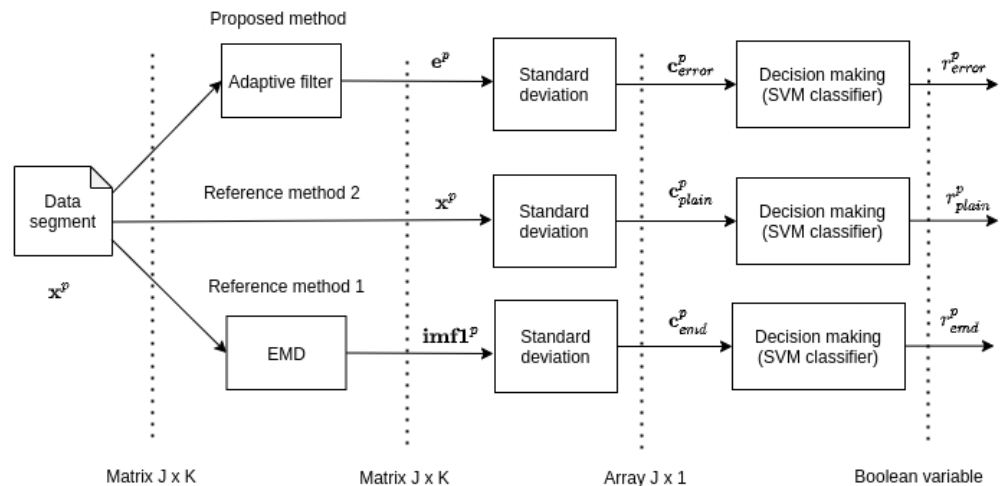


Figure 7. Experimental design—data flow and shapes in the experimental framework.

3. Results

The full cross-validation results are shown in Table 1. Generally, the best results were achieved by the proposed method. The second best method was the one based on EMD filtration, and the worst results were achieved via evaluation of the raw data (denoted as PLAIN in the table). An interesting exception was the combination of channels 3 and 4, where the PLAIN method outperformed the EMD based method. An illustrative example of the extracted features is in Figure 8.

Furthermore, a method’s consistency, shown in Table 1, can be used to estimate the amount of fault-related information in each sensor’s time series. Sensor 1 provided the greatest amount of useful information and dominated the prediction accuracy, and sensor 4 provided the least. To clarify this observation, box plots of the accuracies achieved from sensor 1 (Figure 9) and sensor 4 are shown (Figure 10).

Moreover, note that a higher number of sensors increases the accuracy.

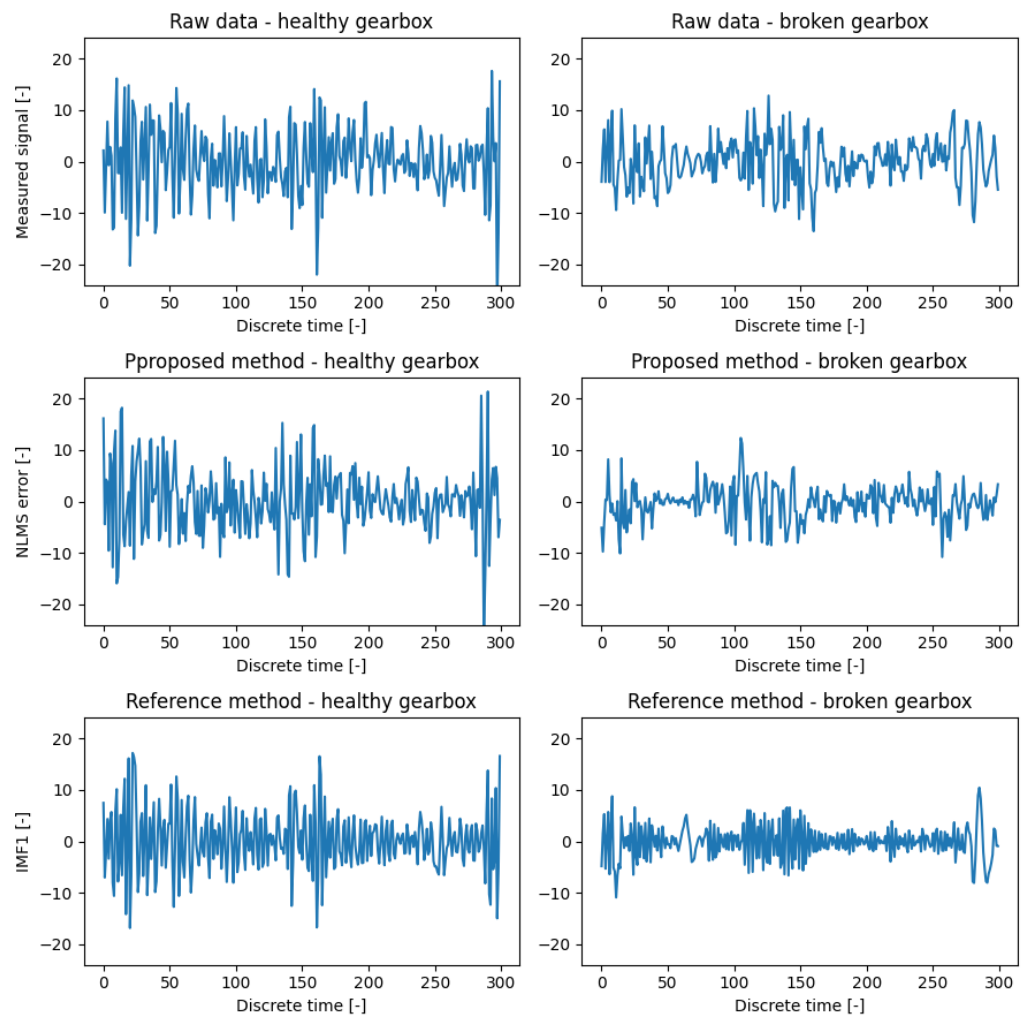


Figure 8. Comparison of features used for proposed and reference methods.

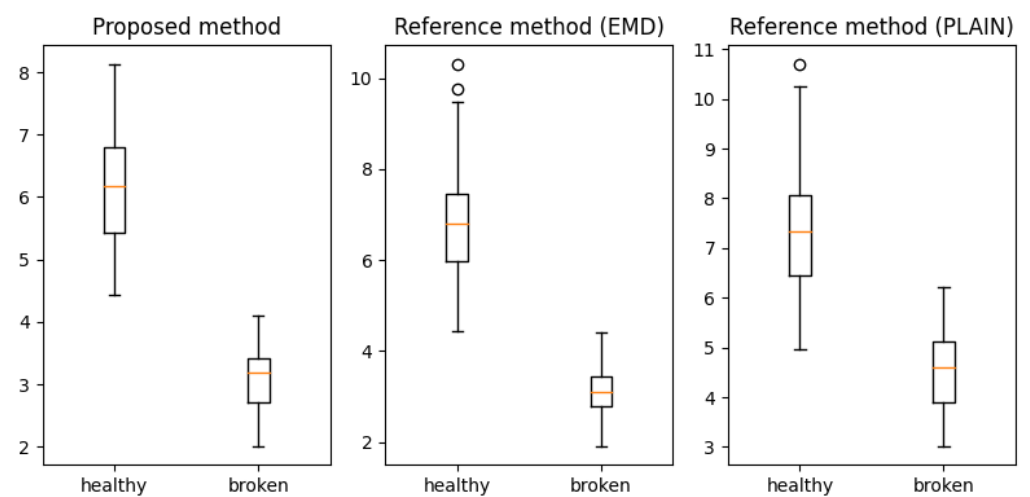


Figure 9. Boxplots of fault descriptors extracted by proposed and reference methods for sensor 1 (the most information).

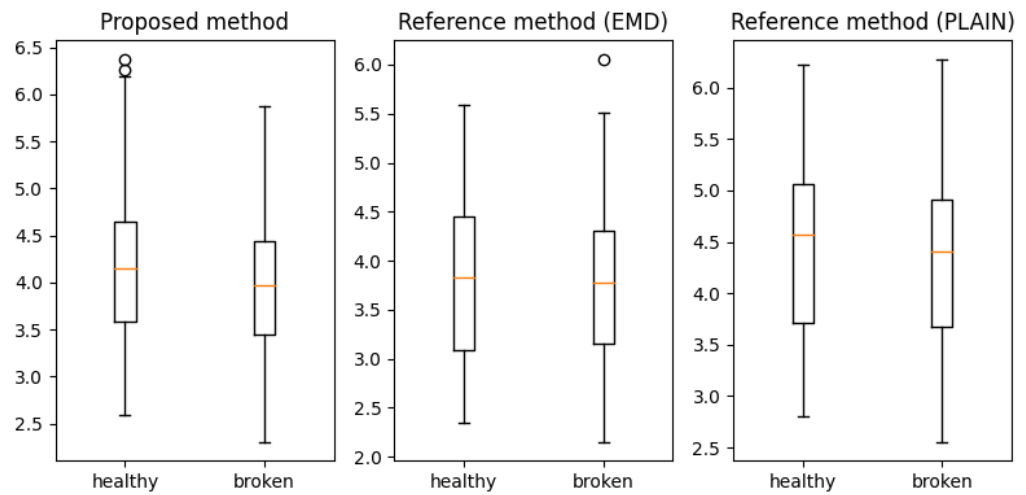


Figure 10. Boxplots of fault descriptors extracted by proposed and reference methods for sensor 4 (the least information).

Table 1. Resulting accuracies of the methods; 100% means the best accuracy, and 50% is equal to classification by a fair coin-flip.

Sensors (SVM Inputs)	Classification Accuracy [%]		
	Proposed Method (NLMS Error)	Reference Method (EMD—IMF 1)	Reference Method (PLAIN)
1, 2, 3, 4	100.0	100.0	100.0
2, 3, 4	93.58	71.534	72.727
3, 4	71.875	58.58	71.875
4	55.114	50.852	52.33
3	65.398	52.67	58.92
2	81.818	71.25	54.886
1	100.0	99.886	94.886

4. Discussion

A novel and time-effective method of gearbox fault detection was proposed in this paper. The proposed method is superior in comparison to the reference methods: raw data statistics and frequency analysis-based feature extraction. Furthermore, the proposed method does not need any information about the measurement setup or the gearbox. The fault detection is based purely on machine learning. The classification criteria are determined according to the data. In other words, the proposed approach can adapt the gearbox diagnosis criteria according to the particular gearbox without human expert input or supervision.

The ultimate goal of this paper was to show that the adaptive filtration error can enhance the information in raw data (the first reference method) in a greater way than the other often-used method of feature extraction—EMD (the second reference method). Therefore, we used a straightforward, noncomplex approach of standard deviation estimation and SVM classification for simple and robust evaluation of the methods. More complex evaluation methods can lead to misinterpreting the feature extraction methods.

The possible utilization of a more advanced classifier (such as neural network) would be a suitable future direction for the research of the proposed feature extraction approach. Another attractive future study might be on applicability research regarding different fault detection tasks, including rotating machinery.

Author Contributions: Conceptualization, J.V. and M.C.; methodology, M.C. and J.V.; software, M.C.; validation, M.C.; formal analysis, J.V. and J.S.; investigation, J.S. and J.V.; resources, J.V.; data curation, J.V.; writing—original draft preparation, J.V., J.S. and Z.K.; writing—review and editing, J.V., J.S., M.C. and Z.K.; visualization, M.C.; supervision, J.V.; project administration, J.V.; funding acquisition, J.V., Z.K. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by a grant of Specific university research—grant IGA A1_FCHI_2021_001. This research received external funding from the ESIF, EU Operational Programme Research, Development and Education, and from the Center of Advanced Aerospace Technology (CZ.02.1.01/0.0/0.0/16_019/0000826), Faculty of Mechanical Engineering, Czech Technical University in Prague.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The dataset we used is publicly available at the address: <https://openi.org/datasets/dataset/gearbox-fault-diagnosis-data> (accessed on 20 August 2021).

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

ASF	adaptive Schur filter
CNN	convolutional neural network
EMD	empirical mode decomposition
FrFT	Fractional Fourier Transform
FT	Fourier transform
HHT	Hilbert–Huang transform
IMF	intrinsic mode function
LMS	least mean squares
LSTM	long short-term memory
NLMS	Normalized least mean squares
SVM	support-vector machine

References

1. Bracewell, R.N.; Bracewell, R.N. *The Fourier Transform and Its Applications*; McGraw-Hill: New York, NY, USA, 1986; Volume 31999.
2. Betta, G.; Liguori, C.; Paolillo, A.; Pietrosanto, A. A DSP-based FFT-analyzer for the fault diagnosis of rotating machine based on vibration analysis. *IEEE Trans. Instrum. Meas.* **2002**, *51*, 1316–1322. [[CrossRef](#)]
3. Patil, S.; Gaikwad, J. Vibration analysis of electrical rotating machines using FFT: A method of predictive maintenance. In Proceedings of the IEEE 2013 Fourth International Conference on Computing, Communications and Networking Technologies (ICCCNT), Tiruchengode, India, 4–6 July 2013; pp. 1–6.
4. Wang, J.; Li, S.; Xin, Y.; An, Z. Gear fault intelligent diagnosis based on frequency-domain feature extraction. *J. Vib. Eng. Technol.* **2019**, *7*, 159–166. [[CrossRef](#)]
5. Luo, J.; Yu, D.; Liang, M. Application of multi-scale chirplet path pursuit and fractional Fourier transform for gear fault detection in speed up and speed-down processes. *J. Sound Vib.* **2012**, *331*, 4971–4986. [[CrossRef](#)]
6. Kar, C.; Mohanty, A. Vibration and current transient monitoring for gearbox fault detection using multiresolution Fourier transform. *J. Sound Vib.* **2008**, *311*, 109–132. [[CrossRef](#)]
7. Vernekar, K.; Kumar, H.; Gangadharan, K. Gear fault detection using vibration analysis and continuous wavelet transform. *Procedia Mater. Sci.* **2014**, *5*, 1846–1852. [[CrossRef](#)]
8. Wang, W.; Wong, A.K. Autoregressive model-based gear fault diagnosis. *J. Vib. Acoust.* **2002**, *124*, 172–179. [[CrossRef](#)]
9. Yang, M.; Makis, V. ARX model-based gearbox fault detection and localization under varying load conditions. *J. Sound Vib.* **2010**, *329*, 5209–5221. [[CrossRef](#)]
10. Heyns, T.; Heyns, P.S.; Zimroz, R. Combining discrepancy analysis with sensorless signal resampling for condition monitoring of rotating machines under fluctuating operations. In Proceedings of the 9th International Conference on Condition Monitoring and Machinery Failure Prevention Technologies 2012 (CM 2012 and MFPT 2012), London, UK, 12–14 June 2012; Volume 1, pp. 503–516. [[CrossRef](#)]
11. Rilling, G.; Flandrin, P.; Goncalves, P. On empirical mode decomposition and its algorithms. In Proceedings of the IEEE-EURASIP Workshop on Nonlinear Signal and Image Processing, Grado, Italy, 3–6 June 2003; Volume 3, pp. 8–11.

12. Dybała, J.; Zimroz, R. Empirical mode decomposition of vibration signal for detection of local disturbances in planetary gearbox used in heavy machinery system. *Key Eng. Mater.* **2014**, *588*, 109–116. [[CrossRef](#)]
13. Guo, T.; Deng, Z. An improved EMD method based on the multi-objective optimization and its application to fault feature extraction of rolling bearing. *Appl. Acoust.* **2017**, *127*, 46–62. [[CrossRef](#)]
14. Liu, B.; Riemenschneider, S.; Xu, Y. Gearbox fault diagnosis using empirical mode decomposition and Hilbert spectrum. *Mech. Syst. Signal Process.* **2006**, *20*, 718–734. [[CrossRef](#)]
15. Wang, Y.H.; Yeh, C.H.; Young, H.W.V.; Hu, K.; Lo, M.T. On the computational complexity of the empirical mode decomposition algorithm. *Phys. A Stat. Mech. Appl.* **2014**, *400*, 159–167. [[CrossRef](#)]
16. Elasha, F.; Ruiz-Carcel, C.; Mba, D.; Chandra, P. A Comparative Study of the Effectiveness of Adaptive Filter Algorithms, Spectral Kurtosis and Linear Prediction in Detection of a Naturally Degraded Bearing in a Gearbox. *J. Fail. Anal. Prev.* **2014**, *14*, 623–636. [[CrossRef](#)]
17. Chaturved, G.K.; Thomas, D.W. Adaptive noise cancelling and condition monitoring. *J. Sound Vib.* **1981**, *76*, 391–405. [[CrossRef](#)]
18. Elasha, F.; Mba, D.; Ruiz-Carcel, C. A comparative study of adaptive filters in detecting a naturally degraded bearing within a gearbox. *Case Stud. Mech. Syst. Signal Process.* **2016**, *3*, 1–8. [[CrossRef](#)]
19. Lee, S.K.; White, P.R. The enhancement of impulsive noise and vibration signals for fault detection in rotating and reciprocating machinery. *J. Sound Vib.* **1998**, *217*, 485–505. [[CrossRef](#)]
20. Makowski, R.; Zimroz, R. Application of Schur Filtering for Local Damage Detection in Gearboxes. *Cond. Monit. Mach. Non-Station. Oper.* **2012**, 301–308. [[CrossRef](#)]
21. Makowski, R.A.; Zimroz, R. Adaptive bearings vibration modelling for diagnosis. In *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*; Springer: Berlin/Heidelberg, Germany, 2011; Volume 6943, pp. 248–259. [[CrossRef](#)]
22. Makowski, R.; Zimroz, R. A procedure for weighted summation of the derivatives of reflection coefficients in adaptive Schur filter with application to fault detection in rolling element bearings. *Mech. Syst. Signal Process.* **2013**, *38*, 65–77. [[CrossRef](#)]
23. Makowski, R.; Zimroz, R. New techniques of local damage detection in machinery based on stochastic modelling using adaptive Schur filter. *Appl. Acoust.* **2014**, *77*, 130–137. [[CrossRef](#)]
24. Yao, X.; Guo, C.; Zhong, M.; Li, Y.; Shan, G.; Zhang, Y. Wind turbine gearbox fault diagnosis using adaptive Morlet wavelet spectrum. In *Proceedings of the 2009 2nd International Conference on Intelligent Computing Technology and Automation, ICICTA 2009, Changsha, China, 10–11 October 2009*; Volume 2, pp. 580–583. [[CrossRef](#)]
25. Nguyen, C.D.; Prosvirin, A.; Kim, J.M. A Reliable Fault Diagnosis Method for a Gearbox System with Varying Rotational Speeds. *Sensors* **2020**, *20*, 3105. [[CrossRef](#)]
26. Tian, S.; Qian, Z. Planetary gearbox fault feature enhancement based on combined adaptive filter method. *Adv. Mech. Eng.* **2015**, *7*, 168781401562032. [[CrossRef](#)]
27. Kucharczyk, D.; Wyłomańska, A.; Zimroz, R. Structural break detection method based on the Adaptive Regression Splines technique. *Phys. A Stat. Mech. Appl.* **2017**, *471*, 499–511. [[CrossRef](#)]
28. Vib, J.; Eng, T.; Elasha, F.; Li, X.; Mba, D.; Ogundare, A.; Ojolo, S. A Novel Condition Indicator for Bearing Fault Detection Within Helicopter Transmission. *J. Vib. Eng. Technol.* **2020**. [[CrossRef](#)]
29. Michalak, A.; Wodecki, J.; Wyłomańska, A.; Zimroz, R. Application of cointegration to vibration signal for local damage detection in gearboxes. *Appl. Acoust.* **2019**, *144*, 4–10. [[CrossRef](#)]
30. Cui, L.; Zhang, Y.; Zhang, F.; Zhang, J.; Lee, S. Vibration response mechanism of faulty outer race rolling element bearings for quantitative analysis. *J. Sound Vib.* **2016**, *364*, 67–76. [[CrossRef](#)]
31. Xueyi, L.; Jialin, L.; Yongzhi, Q.; David, H. Semi-supervised gear fault diagnosis using raw vibration signal based on deep learning. *Chin. J. Aeronaut.* **2020**, *33*, 418–426.
32. Saufi, S.R.; Ahmad, Z.A.B.; Leong, M.S.; Lim, M.H. Gearbox fault diagnosis using a deep learning model with limited data sample. *IEEE Trans. Ind. Inform.* **2020**, *16*, 6263–6271. [[CrossRef](#)]
33. Li, C.; Sanchez, R.V.; Zurita, G.; Cerrada, M.; Cabrera, D.; Vásquez, R.E. Multimodal deep support vector classification with homologous features and its application to gearbox fault diagnosis. *Neurocomputing* **2015**, *168*, 119–127. [[CrossRef](#)]
34. Sun, G.D.; Wang, Y.R.; Sun, C.F.; Jin, Q. Intelligent detection of a planetary gearbox composite fault based on adaptive separation and deep learning. *Sensors* **2019**, *19*, 5222. [[CrossRef](#)] [[PubMed](#)]
35. Heydarzadeh, M.; Kia, S.H.; Nourani, M.; Henao, H.; Capolino, G.A. Gear fault diagnosis using discrete wavelet transform and deep neural networks. In *Proceedings of the IECON 2016—42nd Annual Conference of the IEEE Industrial Electronics Society, Florence, Italy, 24–27 October 2016*; pp. 1494–1500.
36. Chen, Z.; Li, C.; Sanchez, R.V. Gearbox fault identification and classification with convolutional neural networks. *Shock Vib.* **2015**, *2015*, 390134. [[CrossRef](#)]
37. Yao, Y.; Zhang, S.; Yang, S.; Gui, G. Learning attention representation with a multi-scale CNN for gear fault diagnosis under different working conditions. *Sensors* **2020**, *20*, 1233. [[CrossRef](#)]
38. Li, Y.; Cheng, G.; Pang, Y.; Kuai, M. Planetary gear fault diagnosis via feature image extraction based on multi central frequencies and vibration signal frequency Spectrum. *Sensors* **2018**, *18*, 1735. [[CrossRef](#)]
39. Li, C.; Sanchez, R.V.; Zurita, G.; Cerrada, M.; Cabrera, D.; Vásquez, R.E. Gearbox fault diagnosis based on deep random forest fusion of acoustic and vibratory signals. *Mech. Syst. Signal Process.* **2016**, *76*, 283–293. [[CrossRef](#)]

40. Huang, N.E.; Shen, Z.; Long, S.R.; Wu, M.C.; Shih, H.H.; Zheng, Q.; Yen, N.C.; Tung, C.C.; Liu, H.H. The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. *Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci.* **1998**, *454*, 903–995. [[CrossRef](#)]
41. Ha, J.M.; Youn, B.D. A Health Data Map-Based Ensemble of Deep Domain Adaptation under Inhomogeneous Operating Conditions for Fault Diagnosis of a Planetary Gearbox. *IEEE Access* **2021**, *9*, 79118–79127. [[CrossRef](#)]
42. Wang, H.; Xu, J.; Sun, C.; Yan, R.; Chen, X. Intelligent Fault Diagnosis for Planetary Gearbox Using Time-Frequency Representation and Deep Reinforcement Learning. *IEEE/ASME Trans. Mechatron.* **2021**. [[CrossRef](#)]
43. Shi, J.; Peng, D.; Peng, Z.; Zhang, Z.; Goebel, K.; Wu, D. Planetary gearbox fault diagnosis using bidirectional-convolutional LSTM networks. *Mech. Syst. Signal Process.* **2022**, *162*, 107996. [[CrossRef](#)]
44. Azamfar, M.; Singh, J.; Li, X.; Lee, J. Cross-domain gearbox diagnostics under variable working conditions with deep convolutional transfer learning. *J. Vib. Control* **2021**, *27*, 854–864. [[CrossRef](#)]
45. Ye, Z.; Yu, J. Deep morphological convolutional network for feature learning of vibration signals and its applications to gearbox fault diagnosis. *Mech. Syst. Signal Process.* **2021**, *161*, 107984. [[CrossRef](#)]
46. Yu, J.; Zhou, X.; Lu, L.; Zhao, Z. Multiscale Dynamic Fusion Global Sparse Network for Gearbox Fault Diagnosis. *IEEE Trans. Instrum. Meas.* **2021**, *70*, 1–11.
47. Ye, Z.; Yu, J. AKRNet: A novel convolutional neural network with attentive kernel residual learning for feature learning of gearbox vibration signals. *Neurocomputing* **2021**, *447*, 23–37. [[CrossRef](#)]
48. Zhang, K.; Tang, B.; Deng, L.; Liu, X. A hybrid attention improved ResNet based fault diagnosis method of wind turbines gearbox. *Measurement* **2021**, *179*, 109491. [[CrossRef](#)]
49. Zhao, M.; Kang, M.; Tang, B.; Pecht, M. Deep residual networks with dynamically weighted wavelet coefficients for fault diagnosis of planetary gearboxes. *IEEE Trans. Ind. Electron.* **2017**, *65*, 4290–4300. [[CrossRef](#)]
50. Chen, S.N.; Liu, F.; Gao, C.X.; Li, J. Gearbox Fault Diagnosis Classification with Empirical Mode Decomposition Based on Improved Long Short-Term Memory. In Proceedings of the 2021 IEEE 6th International Conference on Cloud Computing and Big Data Analytics (ICCCBDA), Chengdu, China, 24–26 April 2021; pp. 568–575.
51. Wang, Y.; Sun, T.; Liu, Y. Gearbox Fault Diagnosis Based on Two-Class NMF Network Under Variable Working Conditions. *J. Electr. Eng. Technol.* **2021**, 1–12. [[CrossRef](#)]
52. Xiao, D.; Ding, J.; Li, X.; Huang, L. Gear fault diagnosis based on kurtosis criterion VMD and SOM neural network. *Appl. Sci.* **2019**, *9*, 5424. [[CrossRef](#)]
53. Aherwar, A.; Khalid, M.S. Vibration analysis techniques for gearbox diagnostic: A review. *Int. J. Adv. Eng. Technol.* **2012**, *3*, 4–12.
54. Sharma, V.; Parey, A. A review of gear fault diagnosis using various condition indicators. *Procedia Eng.* **2016**, *144*, 253–263. [[CrossRef](#)]
55. Basangar, S.; Tripathi, B. Literature review on fault detection of equipment using machine learning techniques. In Proceedings of the IEEE 2020 International Conference on Computation, Automation and Knowledge Management (ICCAKM), Dubai, United Arab Emirates, 9–10 January 2020; pp. 62–67.
56. Malik, H.; Pandya, Y.; Parashar, A.; Sharma, R. Feature extraction using EMD and classifier through artificial neural networks for gearbox fault diagnosis. In *Applications of Artificial Intelligence Techniques in Engineering*; Springer: Berlin/Heidelberg, Germany, 2019; pp. 309–317.
57. Haykin, S.S. *Adaptive Filter Theory*; Pearson Education India: Delhi, India, 2005.
58. Huang, N.E. *Hilbert-Huang Transform and Its Applications*; World Scientific: Singapore, 2014; Volume 16,
59. Cortes, C.; Vapnik, V. Support-vector networks. *Mach. Learn.* **1995**, *20*, 273–297. [[CrossRef](#)]
60. Patle, A.; Chouhan, D.S. SVM kernel functions for classification. In Proceedings of the IEEE 2013 International Conference on Advances in Technology and Engineering (ICATE), Mumbai, India, 23–25 January 2013; pp. 1–9.