


## Article

# Effective Fault Detection and Diagnosis for Power Converters in Wind Turbine Systems Using KPCA-Based BiLSTM

Zahra Yahyaoui <sup>1</sup>, Mansour Hajji <sup>1</sup>, Majdi Mansouri <sup>2,3,\*</sup>, Kamaleldin Abodayeh <sup>3</sup> , Kais Bouzrara <sup>4</sup> and Hazem Nounou <sup>2</sup>

<sup>1</sup> Research Unit Advanced Materials and Nanotechnologies (UR16ES03), Higher Institute of Applied Sciences and Technology of Kasserine, Kairouan University, Kasserine 1200, Tunisia

<sup>2</sup> Electrical and Computer Engineering Program, Texas A&M University at Qatar, Doha P.O. Box 23874, Qatar

<sup>3</sup> Department of Mathematical Sciences, Prince Sultan University, Riyadh 12435, Saudi Arabia

<sup>4</sup> Laboratory of Automatic Signal and Image Processing, National Engineering School of Monastir, Monastir 5035, Tunisia

\* Correspondence: majdi.mansouri@qatar.tamu.edu

**Abstract:** The current work presents an effective fault detection and diagnosis (FDD) technique in wind energy converter (WEC) systems. The proposed FDD framework merges the benefits of kernel principal component analysis (KPCA) model and the bidirectional long short-term memory (BiLSTM) classifier. In the developed FDD approach, the KPCA model is applied to extract and select the most effective features, while the BiLSTM is utilized for classification purposes. The developed KPCA-based BiLSTM approach involves two main steps: feature extraction and selection, and fault classification. The KPCA model is developed in order to select and extract the most efficient features and the final features are fed to the BiLSTM to distinguish between different working modes. Different simulation scenarios are considered in this study in order to show the robustness and performance of the developed technique when compared to the conventional FDD methods. To evaluate the effectiveness of the proposed KPCA-based BiLSTM approach, we utilize data obtained from a healthy WTC, which are then injected with several fault scenarios: simple fault generator-side, simple fault grid-side, multiple fault generator-side, multiple fault grid-side, and mixed fault on both sides. The diagnosis performance is analyzed in terms of accuracy, recall, precision, and computation time. Furthermore, the efficiency of fault diagnosis is shown by the classification accuracy parameter. The experimental results show the efficiency of the developed KPCA-based BiLSTM technique compared to the classical FDD techniques (an accuracy of 97.30%).

**Keywords:** wind energy conversion (WEC); fault detection and diagnosis (FDD); kernel PCA (KPCA); bidirectional long short term memory (BiLSTM)



**Citation:** Yahyaoui, Z.; Hajji, M.; Mansouri, M.; Abodayeh, K.; Bouzrara, K.; Nounou, H. Effective Fault Detection and Diagnosis for Power Converters in Wind Turbine Systems Using KPCA-Based BiLSTM. *Energies* **2022**, *15*, 6127. <https://doi.org/10.3390/en15176127>

Academic Editors: Davide Astolfi and Abu-Siada Ahmed

Received: 29 March 2022

Accepted: 31 May 2022

Published: 23 August 2022

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



**Copyright:** © 2022 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

Wind energy is one of most essential substitute energies due its competitive cost and maturity of technology. According to the World Wind Energy Association (WWEA), the total capacity of all wind farms worldwide reached 744 GW in 2020.

Due to the development of wind power production, enhancement of the control of wind energy conversion (WEC) systems is required. For this reason, manufacturers' efforts have been focused on the improvement of these systems' lifetimes and the decrease of operation breakdowns (downtime maintenance process), leading to continuous energy production with high power quality [1,2].

Wind energy conversion (WEC) systems are composed of various interconnected electrical and mechanical elements. However, unexpected failures usually accompany the operation of these systems. When a fault in a system occurs, it can have an adverse effect on the system's availability, in addition to the production rate. Indeed, many components of wind turbines (WT) can fail due to harsh environmental and operating conditions, resulting

in lengthy downtime maintenance periods [3,4]. The most common failures are related to blades [5,6], generators [7,8], power converters [1,9], and gearboxes [10,11]. As a crucial component and the heart of these systems, the power converter plays a significant role in transferring the generated power to the grid. It converts electrical energy that varies according to the wind speed to energy with a constant frequency complying with grid specifications [12]. It was indicated in [13] that 21% of 25% of the total failures in WEC converters (WECC) are caused by the semiconductor. In order to avoid the WECC collapse, these failures should be detected and diagnosed at an early stage. Therefore, fault detection and diagnosis (FDD) is viewed as essential means to achieve these goals [14]. The authors of [15,16] considered multiple faults in the same-side converter. They address multiple faults in both converter sides at once. The authors of [17] have studied multiple faults by modeling both converter sides as a state space equation. In [18], the authors examined two open-switch faults in one sub-module and also addressed the detection of multiple faults in random sub-module elements. However, the linking effects between generator-side and grid-side converters are not taken into account, which could affect considerably the system behavior. The authors in [14] focused on simple faults in both converter sides. This current work deals with faults in both converter sides, taking into consideration all possible fault scenarios such as simple fault generator-side, simple fault grid-side, multiple fault generator-side, multiple fault grid-side, and mix fault on both sides. Each scenario affects the system behavior in a different way, accordingly, considering each of it is a crucial task.

Generally, FDD approaches can be categorized into two main classes: the model-based and the data-driven methods. Model-based FDD uses observers and system identification models of the processes; it demands a precise mathematical model, which is complicated to acquire in reality. Its performance is dramatically impacted by uncertainties and unmodeled noises [19,20]. Data-driven methods aim to extract information from the measured signals to train the model, and then use the information for diagnosis in the testing phase [21–23]. Numerous studies based on machine learning approaches have been employed in WEC FDD, such as decision tree (DT) [24], naive Bayes (NB) [25], support vector machine (SVM) [26], K-nearest neighbors (KNN) [27], and random forest (RF) [14]. In [2], a WEC fault diagnosis technique based on an RF and kernel principal component analysis (KPCA) approach is developed. In this proposal, KPCA is applied to extract the most informative features from data, with the aim of improving the classification results using an RF classifier. In [24], the authors introduce five-stage statistical process control and machine learning methods to diagnose wind turbine faults (rotary blades, gearboxes, generators, and hydraulic oil systems) and predict maintenance demands. The five adopted analytical tools in statistical process control are: (1) check lists, (2) Pareto charts, (3) cause and effect diagrams, (4) scatter plots, and (5) control charts. Firstly, the check list comprises information such as the type of wind turbine faults, the duration of faults, causes, and repair events. Authors have classified the repair events by frequency of anomalies in the dataset. Secondly, a Pareto chart is developed based on the classified check list items and presents the repair events with regard to cumulative percentage. Thirdly, an analytical tool, that is, the cause and effect diagram, is presented in order to distinguish the essential causes of principal mechanical issues and produce recommendations to technicians for maintenance. Fourthly, scatter plots are applied to investigate the relationship between features and determine abnormal data. Lastly, control charts are applied to show changes and variation in the observed data over time. After that, a density-based spatial clustering of applications with noise (DBSCAN) approach is used to represent the relationship between the entire amount of wind generation and the five attributes, in addition to ranking normal and abnormal data. Finally, two machine learning techniques—decision tree and random forest—are applied in order to construct a predictive maintenance models for anomalies. The inherent disadvantages of traditional ML-based approaches make them ineffective at representing complex functions due to their unsatisfactory performance and their generalization capabilities. With the explosion of deep learning (DL) algorithms in artificial intelligence (AI) applications, technology has shown a strong ability to surpass

conventional intelligent algorithms [28], whose problems include their dependence on hand-designed feature, as well as their difficulty in understanding sequential data. Thus, many researchers have opted to use DL modes instead of traditional classifiers in fault diagnosis. In fact, the major distinction between AI models and DL models is that the latter can automatically learn precious features directly from raw data. Considering the rapid rise of DL, many architecture have been developed, such as convolutional neural networks (CNN), deep belief networks (DBN), and recurrent neural networks (RNN).

Authors in [29] propose an ensemble transfer CNN driven by multi-channel signals for fault diagnosis of rotating machinery. In this case, modified CNNs based on stochastic pooling and leaky rectified linear unit (LReLU) are pre-trained using multi-channel signals. Then, the target CNN is initialized using the learned parameter knowledge of each individual source CNN with the help of parameter transfer. Lastly, in order to achieve the comprehensive result, a new decision fusion procedure is constructed to flexibly fuse each individual target CNN. An FDD approach based on the convolutional neural network long short-term memory attention mechanism (CNN-LSTM-AM) for anomaly recognition and fault detecting of wind turbine is suggested in [30]. The CNN is used to extract features of state space from wind turbine, LSTM is applied to improve the time characteristics fusion of different part states, and AM is used to help the model make more accurate judgments through mapping weight and parameter learning. The authors of [31] propose an approach to regularize the discriminant structure of the deep network with both intrinsic and extrinsic generalization goals in order to improve the learning of robustness features and to generalize to unseen domains. In [32], the authors develop an improved RNN techniques for fault detection and diagnosis for wind energy conversion (WEC) systems. In the beginning, a reduced RNN-based hierarchical K-means clustering is adopted in order to simplify the complexity of the model in terms of training and computation time. It is used to treat the correlations between samples and extract a reduced number of observations from the training data matrix. Then, two reduced RNN-based interval-valued-data methods are developed for classification purposes.

With the RNN, sequence inputs of variable length can be handled due to the recurrent hidden states, whose activation at any particular time is dependent on that of the previous moment. Other research proposes long short-term memory (LSTM) to directly learn features and time-series data [33]. In fact, the recursive behavior of the LSTM gate architecture allows it to capture long-term dependencies and efficiency figures without the gradient vanishing problem of recurrent neural networks (RNNs) [34].

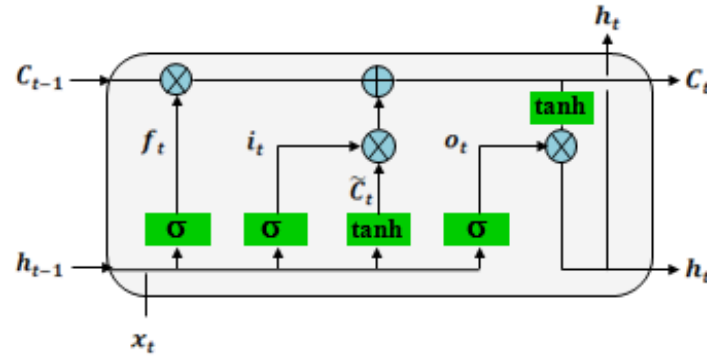
In the current work, we propose an innovative fault diagnosis paradigm using KPCA-based BiLSTM. In fact, the previous studied LSTM-based fault diagnosis approaches were applied directly to raw data without taking account the impact on the extracted and selected features on the classification accuracy, as well as the nonlinear behavior of features. To address these issues, a KPCA-based bidirectional LSTM (KPCA-based BiLSTM) FDD approach is proposed to detect the faults and distinguish between the working modes in the WTC systems. The KPCA model is able to deal with noisy, nonlinear, multivariate, and statistical features [35]. In comparison to other nonlinear techniques, KPCA has the advantages of not involving nonlinear optimization, requiring no prior specification of reduced space dimensions, and being able to handle a wide range of nonlinearities due to its ability to use different kernels [36]. Therefore, in this work, the KPCA feature extraction/selection paradigm and the BiLSTM classification model are applied to detect and classify the WTC faults. The proposed approach makes full use of the KPCA for powerful feature extraction/selection and BiLSTM for fault diagnosis, which can solve the problem of nonlinear, statistical, and multivariate feature extraction and fault diagnosis in WTC systems.

This paper is organized as follows: Section 2 is dedicated to a brief description of the KPCA tool used in feature extraction and selection and of the BiLSTM technique for classification purposes. Section 3 presents the application of the developed methodology for fault detection and diagnosis. Finally, the conclusions are illustrated in Section 4.

## 2. Model

### 2.1. Bidirectional LSTM Description

LSTM was derived from recurrent neural networks (RNN) in 1997 by Hochreiter and Schmidhuber [28]. It was developed to tackle the vanishing gradient issue witnessed in RNNs. Hence, to achieve this target, the architecture of the LSTM has three gates: the input gate, forget gate, and output gate. Figure 1 illustrates the LSTM cell with input gate ( $i_t$ ), forget gate ( $f_t$ ), and output gate ( $o_t$ ), which are denoted by the following equations:



**Figure 1.** Architecture of LSTM.

The forget gate ( $f_t$ ) indicates what information of the previous state ( $C_{t-1}$ ) will be forgot or kept by looking at the values of the current input vector ( $x_t$ ) and hidden state ( $h_{t-1}$ ), as given in the following equation:

$$f_t = \sigma(W_f[h_{t-1}, x_t] + b_f) \quad (1)$$

where  $W_f$  and  $b_f$  represent the weight matrix and the bias term, respectively.

In the same way and in order to update the cell state, the input gate ( $i_t$ ) decides how much information from the input ( $x_t$ ) and ( $h_{t-1}$ ) must pass, expressed as:

$$i_t = \sigma(W_i[h_{t-1}, x_t] + b_i) \quad (2)$$

$$\tilde{C}_t = \tanh(W_c[h_{t-1}, x_t] + b_c) \quad (3)$$

where  $\tilde{C}_t$  denotes an immediate condition.

The updated state of the cell (when deciding which information to reserve and which to forget) is presented as follows:

$$C_t = f_t \odot C_{t-1} \odot i_t \odot \tilde{C}_t \quad (4)$$

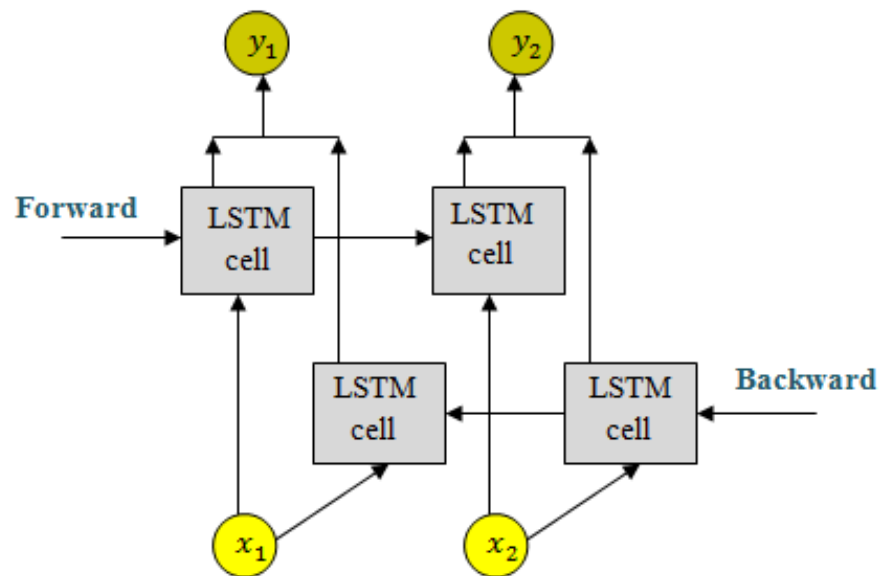
where  $C_t$  represents the long term state and the symbol  $\odot$  denotes element-wise vector multiplication. The output gate ( $o_t$ ) checks the flow of information from the current cell state to the hidden state.

$$o_t = \sigma(W_o[h_{t-1}, x_t] + b_o), \text{ and} \quad (5)$$

$$h_t = o_t \odot \tanh(C_t) \quad (6)$$

where  $h_t$  denotes the output. LSTM exists in several architectures [37] and it might be used in the following forms: vanilla LSTM, stacked LSTM, CNN-LSTM, encoder–decoder LSTM, and bidirectional LSTM. The last of these is the focus of this study.

In 2005, Graves and Schmidhuber developed the bidirectional LSTM by fusing the BRNN with the LSTM cell. The sequential data have strong temporal dependencies in machine disease monitoring systems [38]. Thus, it is important to take into consideration the future situation [39]. Accordingly, the BiLSTM is an essential means of handling this case. Figure 2 illustrates the general concept of BiLSTM architecture.



**Figure 2.** Architecture of BiLSTM.

The architecture for classification purposes is shown in Figure 3. It is composed of an input layer, a BiLSTM layer followed by a fully connected layer, and a softmax layer at the output.

BiLSTM can learn input in both directions: forward and backward. The forward LSTM treats data from left to right and its hidden state can be expressed as  $\vec{h}_t = LSTM(x_t, \vec{h}_{t-1})$ , while the backward LSTM treats information in the opposite direction, and its hidden state can be presented as  $\overleftarrow{h}_t = LSTM(x_t, \overleftarrow{h}_{t+1})$ . Finally, concatenate the forward and backward states to generate the BiLSTM output, as presented in the following equation:

$$h_t = [\vec{h}_t, \overleftarrow{h}_t] \quad (7)$$

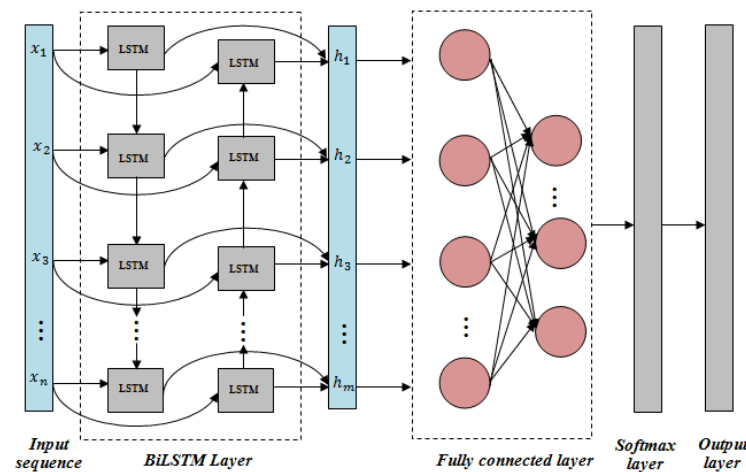
Take into consideration that the final hidden state  $h_f$  encodes the most features from the input signal and uses this as input to the fully connected layer, which aims to convert it into a vector in which the length is equal to the class number. A softmax layer is approved for fault classification. The probability distribution is given as:

$$\tilde{Y} = \text{softmax}(W_s h_f + b_s) \quad (8)$$

where  $W_s$  and  $b_s$  indicate the weight and bias, respectively.

$$\text{softmax}(z_i) = \frac{\exp(z_i)}{\sum_{j=1}^k \exp(z_j)} \quad (9)$$

where  $z_i$  is the  $i$ th element of the input vector  $z$ . The BiLSTM model is trained by minimizing the error between the predicted  $\tilde{Y}$  and actual  $Y$ .



**Figure 3.** BiLSTM architecture for fault classification.

## 2.2. Kernel Principal Component Analysis

KPCA extends conventional PCA to handle nonlinear data [40]. In fact, KPCA includes two main steps: (1) mapping the data into a higher dimensional feature space, and (2) performing the linear PCA in that space.

### 2.2.1. KPCA-Based Feature Extraction

Consider the data matrix  $X = [X_1 \ X_2 \ \cdots \ X_m] \in \mathbb{R}^{N \times m}$ , where  $N$  denotes the number of samples and  $m$  represents the number of variables. The mapped data in the new feature space is organized as follows:

$$X = [\phi(x_1) \ \phi(x_2) \ \cdots \ \phi(x_N)]^T \in \mathbb{R}^{N \times h} \quad (10)$$

where  $h \gg m$  is the dimension of the feature space. The following Eigenvector expression is used to compute the kernel principal components (KPCs):

$$K\alpha = \lambda\alpha \quad (11)$$

This equation indicates that  $\alpha$  and  $\lambda$  are the Eigenvectors and Eigenvalues of the kernel matrix  $K$ . The kernel matrix  $K$  is declared as:

$$K = XX^T = \begin{bmatrix} k(x_1, x_1) & \cdots & k(x_1, x_N) \\ \vdots & \ddots & \vdots \\ k(x_N, x_1) & \cdots & k(x_N, x_N) \end{bmatrix} \quad (12)$$

### 2.2.2. KPCA-Based Features Selection

The feature selection function intends to select the smallest group with most relevant and expressive features by eliminating all the irrelevant and redundant features. The given equation represents the kernel matrix Eigenvector [41]:

$$v = \lambda^{-1} X^T \alpha \quad (13)$$



The matrix  $P = [v_1, \dots, v_\ell]$  denotes the matrix of the  $\ell$  retained principal loading of the KPCA in the feature space. Referring to Equation (2), the matrix  $P$  can be defined as:

$$\begin{aligned} P &= [\frac{1}{\lambda_1} X^T \alpha_1, \dots, \frac{1}{\lambda_\ell} \alpha_\ell] \\ &= \begin{bmatrix} \lambda_1^{-1/2} X^T \alpha_1^*, \dots, \lambda_\ell^{-1/2} X^T \alpha_\ell^* \end{bmatrix} \\ &= X^T P^* \Lambda^{-1/2} \end{aligned} \quad (14)$$

where  $P^* = [\alpha_1^*, \dots, \alpha_\ell^*]$  and  $\Lambda = \text{diag}(\lambda_1, \dots, \lambda_\ell)$  are the  $\ell$  principal Eigenvectors and Eigenvalues of  $K$ , respectively.

Moreover, the kernel principal components are computed as:

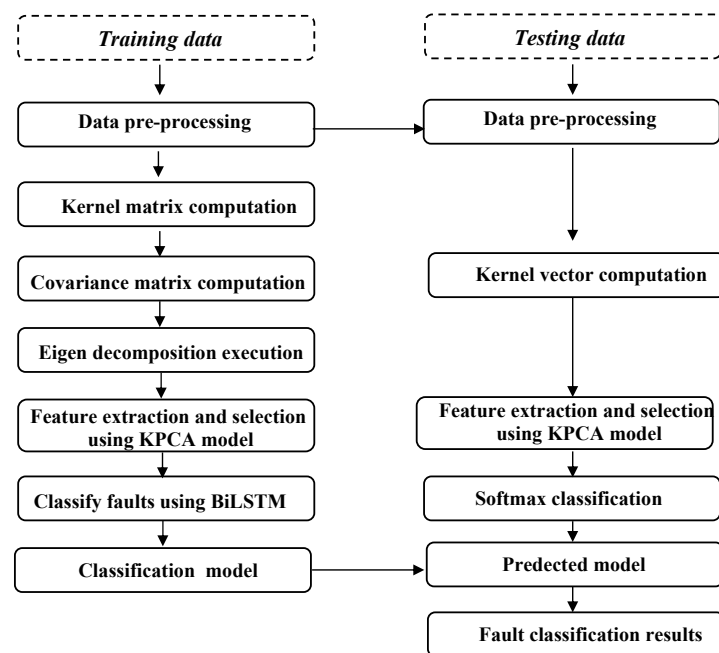
$$t = \Lambda^{-1/2} P^{*T} k(x) \quad (15)$$

The selection of the number of kernel principal components (KPCs) has been subjected to various studies; Ref. [42] details some of them. In this work, the cumulative percent variance (CPV) criterion is used to select the first KPCs in the KPCA model. The features extracted from the KPCA model are the first retained KPCs.

### 3. Proposed Approach and Case Study Experiment

#### 3.1. Proposed Approach

The proposed methodology includes two major steps comprising feature selection and extraction, and fault classification. The developed approach is discussed in such a way that the kernel principal component analysis (KPCA) method is applied for feature extraction goals and the BiLSTM classifier is used for fault diagnosis. The goal of this methodology is to reduce the complexity of the proposed classifier. The first step of the proposed methodology includes the gathering of WECC data. Then, the KPCA is applied to the data in the interest of extracting and selecting the most effective and relevant features. In the next step, the final features subset is considered as input to the BiLSTM tool to classify faults and distinguish between the different operating mode. To summarize, the current paper presents an intelligent fault diagnosis approach based on the KPCA model and the BiLSTM classifier. In fact, the classical BiLSTM-based fault diagnosis techniques were previously utilized directly on the raw data without considering the impact on the feature extraction and selection phase in the diagnosis performance. To deal with these issues, a multivariate KPCA-based bidirectional LSTM classifier approach is presented to detect and identify the faults in WTC systems. In the developed FDD approach (so-called KPCA-based BiLSTM), the KPCA model is applied to extract nonlinear, multivariate, and statistical features, and BiLSTM is utilized for fault classification purposes. The proposed KPCA-based BiLSTM approach makes full application of the KPCA for feature extraction/selection and BiLSTM for fault classification, which can solve the problem of nonlinear, statistical and multivariate feature extraction and fault diagnosis in WTC systems. The flowchart of the developed KPCA-based BiLSTM algorithm is illustrated in Figure 4. Algorithm 1 shows the principal steps of the developed KPCA-based BiLSTM algorithm.



**Figure 4.** Architecture of the proposed approach.

---

**Algorithm 1** KPCA-based BiLSTM Algorithm

---

Input:  $N \times m$  data matrix  $X_i, i = 1, 2, \dots, n$

**Training phase**

1. Standardize the training data set;
2. Determine the KPCA model;
3. Extract and select the more relevant features using the KPCA model;
4. Classify the faults through the BiLSTM classifier;
5. Ascertain the classification model.

**Testing phase**

1. Standardize the testing data set;
  2. Calculate the kernel vector;
  3. Extract and select features using the KPCA model;
  4. Classify the faults through the BiLSTM classifier;
  5. Establish the prediction model;
  6. Achieve the fault diagnosis results.
- 

### 3.2. System Description

In this paper, a variable speed wind turbine based on a squirrel cage induction generator (SCIG) is considered, as illustrated in Figure 5. This structure offers unlimited variable speed operation. No matter the rotation speed of the machine, the voltage created is rectified and converted into direct current and voltage. Accordingly, the grid-side converter command assists in giving an alternating voltage with a constant frequency referring to that of the grid. The maximum power generated by the turbine is determined by the nominal power of the generator. For this configuration, the grid-side generator is based on an insulated gate bipolar transistor (IGBT), the structure of which is the same as that of the grid-side converter. The wind turbine parameters are shown in Table 1.



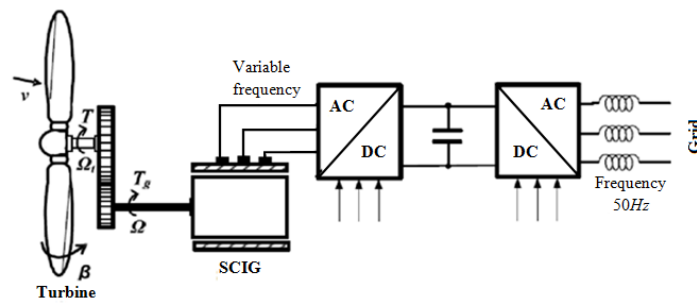


Figure 5. Variable speed wind turbine based on SCIG [14].

Table 1. Wind turbine parameters.

Parameters	Nomenclature	Values
Nominal power of turbine	$P_{tn}$	15 kW
Moment of inertia of turbine	$J_t$	1000 kgm <sup>2</sup>
Stator resistance	$R_s$	0.087 Ohm
Stator leakage inductance	$L_s$	0.8 mH
Rotor resistance	$R_r$	0.228 Ohm
Rotor resistance	$R_r$	0.228 Ohm
Rotor leakage inductance	$L_r$	0.8 mH
Magnetizing inductance	$L_m$	34.7 mH
Number of poles	$P$	4
Moment of inertia of generator	$J_g$	0.2 kgm <sup>2</sup>

The power converters are a crucial component in WEC systems. The authors of [43] proved that 21% of the faults in power converters are attributed to semiconductors (IGBT, diode), as shown in Figure 6.

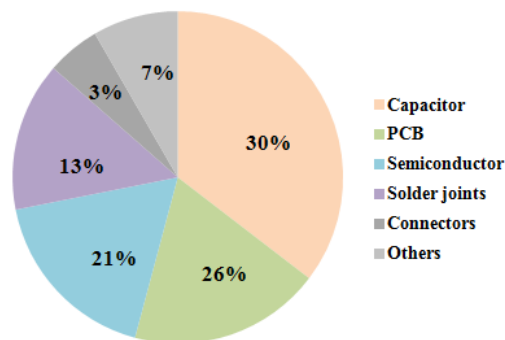


Figure 6. Common catastrophic failures of IGBT.

The usual faults in power switches involve two type of failures: wear-out failures and catastrophic failures. The first type ensues from long time degradation, while catastrophic faults generally happen due to one overstress incident. This paper concern only open-circuits and short-circuits, which cause irretrievable harm to the converter system. In fact, open-circuit faults of IGBT do not cause serious damage to the converter, but influence the performance of the other-side converter and the feedbacks in the control loop. Figure 7 classifies IGBT catastrophic failures into open-circuit and short-circuit statuses arising from various failure mechanisms.

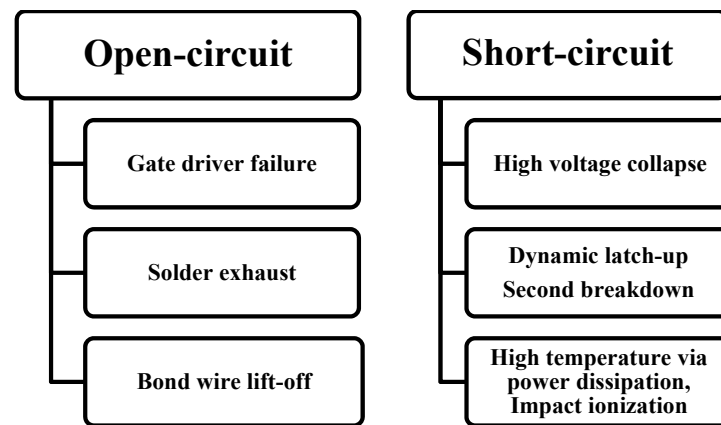


Figure 7. Common catastrophic failures of IGBT.

In the wind chain, the power converter topology exists on two levels. Each converter is composed of three arms. Each arm includes a high and a low IGBT, (as shown in Figure 8).

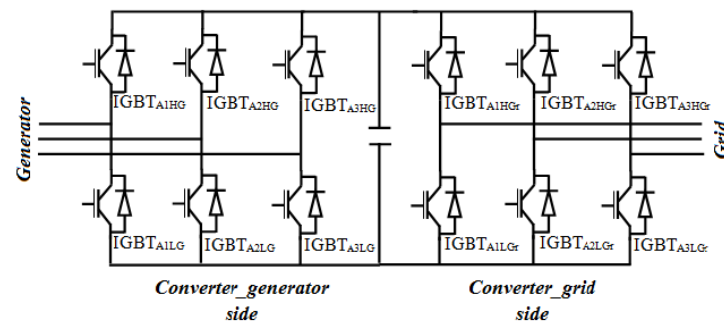
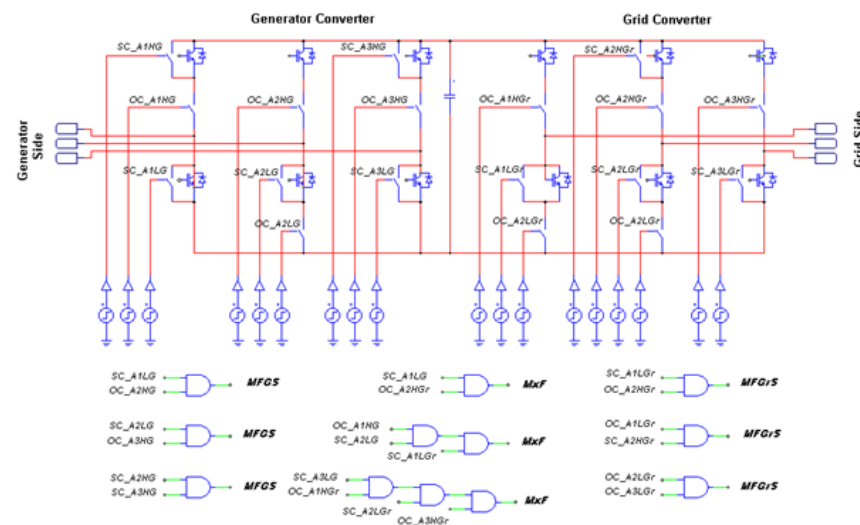


Figure 8. Converter topology [14].

### 3.3. Data Collection

In order to construct a data base to perform FDD, a test bench must be designed under realistic conditions. As detailed in Figure 9, the test setup should be positioned to stress the IGBT modules as they would be in a real wind turbine application. For the sake of injecting short-circuits and open-circuits, we add a controlled switch either in parallel or in serial. This paper deals with several fault scenarios and each scenario comprises different cases, as shown in Table 2.

- **First scenario:** This denotes simple faults that concern just one IGBT on the generator-side converter (SFGS);
- **Second scenario:** This denotes simple faults that concern just one IGBT on the grid-side converter (SFGGrS);
- **Third scenario, forth scenario:** Practically, there can be more than one fault on the same converter side; in this paper, we consider multiple faults on the generator side (MFGS) and grid side (MFGGrS) separately;
- **Fifth scenario:** In the real word, faults may happen on both the converter sides simultaneously; for that reason, we consider mixed faults (MxF);
- **Sixth scenario:** In order to monitor the system in all its states, we combine all the above scenarios.

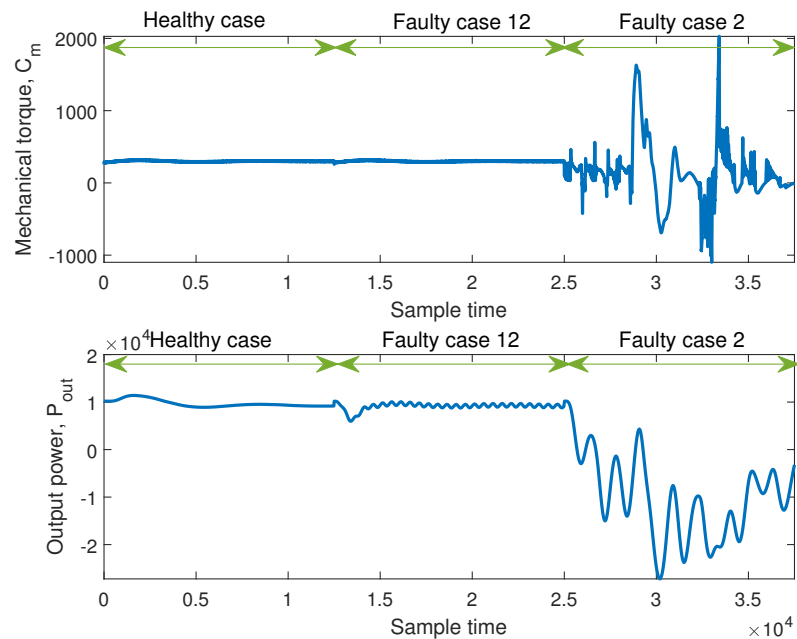


**Figure 9.** Illustration of failure mode distribution.

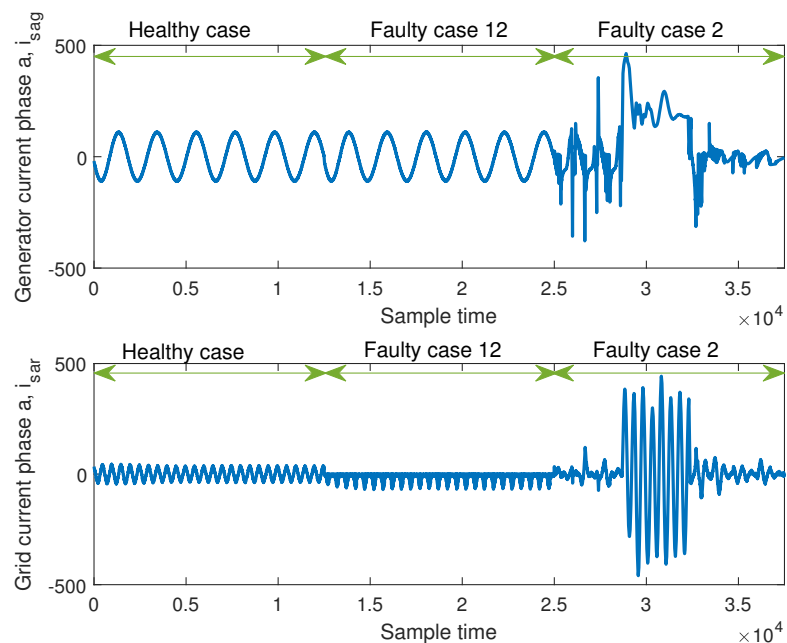
**Table 2.** Construction of database for the fault diagnosis system.

Type of Fault	Symbol	Fault Scenario
Simple fault generator-side	SFGS	Case 1: Short circuit (SC-A1HG) ( $IGBT_{A1HG}$ ). Case 2: Open circuit (OC-A2LG) ( $IGBT_{A2LG}$ ). Case 3: Open circuit (OC-A3HG) ( $IGBT_{A3HG}$ ).
Simple fault grid-side	SFGrS	Case 4: Open circuit (OC-A1LGr) ( $IGBT_{A1LGr}$ ). Case 5: Short circuit (SC-A2HGr) ( $IGBT_{A2HGr}$ ). Case 6: Short circuit (SC-A3LGr) ( $IGBT_{A3LGr}$ ).
Multiple faults generator-side	MFGS	Case 7: Short circuit (SC-A1LG) ( $IGBT_{A1LG}$ ) and Open circuit (OC-A2LG) ( $IGBT_{A2LG}$ ). Case 8: Short circuit (SC-A2LG) and Open circuit (OC-A3HG) ( $IGBT_{A3HG}$ ). Case 9: Short circuit ( $IGBT_{A2HG}$ ) and Short (SC-A2HG and SC-A2HG) ( $IGBT_{A2HG}$ ).
Multiple faults grid-side	MFGrS	Case 10: Short circuit ( $IGBT_{A1LGr}$ ) and Open circuit (SC-A1LGr and OC-A2HGr) ( $IGBT_{A2HGr}$ ). Case 11: Open circuit ( $IGBT_{A1LGr}$ ) and Short circuit (OC-A1LGr and SC-A2HGr) ( $IGBT_{A2HGr}$ ). Case 12: Open circuit ( $IGBT_{A2LGr}$ ) and Open circuit (OC-A2LGr and OC-A3LGr) ( $IGBT_{A3LGr}$ ).
Mixed fault both sides	MxF	Case 13: Short circuit ( $IGBT_{A1LG}$ ) and Open circuit (SC-A1LG and OC-A2HGr) ( $IGBT_{A2HGr}$ ). Case 14: Open circuit ( $IGBT_{A1HG}$ ), Short circuit ( $IGBT_{A2LG}$ ) and Short circuit (OC-A1HG, SC-A1LGr and SC-A2LG) ( $IGBT_{A1LGr}$ ). Case 15: Short circuit ( $IGBT_{A3LG}$ ), Open circuit ( $IGBT_{A1HG}$ ), Short circuit ( $IGBT_{A2LGr}$ ) and Open circuit (SC-A3LG, OC-A3HGr, OC-A1HGr, SC-A2LGr) ( $IGBT_{A3HGr}$ ).

Figures 10 and 11 show the behavior of some electrical and mechanical variables in different faulty cases.



**Figure 10.** Input torque and output power for different cases.



**Figure 11.** Generator current and grid current for different cases.

Figures 10 and 11 clearly demonstrate that faults do not affect the system behavior in the same way. In fact, some fault scenarios do not significantly affect the behavior of the system, in which case service can be maintained until the fault is isolated, as illustrated in Case 12. For example, the output power in healthy mode is almost constant, while when the fault is injected, the same level of power is found with some oscillations. Other types of faults that considerably affect the behavior of the system are considered serious. In Case 2, for instance, the generator current reaches around 500A, which is an insupportable current for the system, and in this situation the system must be taken out of service immediately.

## 4. Results and Discussions

### 4.1. Performance Metrics

In order to evaluate and compare performance, the approved criteria are: accuracy (%), which indicates the rate of observations correctly predicted over the total number of observations; recall (%), which indicates, in the pertinent class, the rate of positive observations correctly predicted to observations; precision (%), which indicates the number of positive observations correctly predicted divided by the number of total predicted positive observations; F1 score (%), which indicates the weighted average of precision and recall; and computation time (CT(s)), which represents the time required to carry out the algorithm.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (16)$$

$$Recall = \frac{TP}{TP + FN} \quad (17)$$

$$Precision = \frac{TP}{TP + FP} \quad (18)$$

$$F_1Score = 2 \times \frac{precision \times recall}{precision + recall} \quad (19)$$

where *TP* (true positive) is properly classified positive observations, *FP* (false positive) is mis-classified positive observations, *TN* (true negative) is correctly classified negative samples, and *FN* (false negative) is misclassified negative observations.

### 4.2. Parameters Setting

In this work, the 95% cumulative variance criterion is applied to select the retained KPCs where 32 KPCs are maintained. Sampling noise can appear during the training process due to the complex relationships among inputs and outputs of neural networks, leading to overfitting, which decreases the predictive capability of the model [44]. In order to avoid this issue, optimal hyperparameters are used in this paper (as shown in Table 3), such as the Adam optimization algorithm, which is used in order to decrease the error in each iteration. Actually, Adam exceeds other optimization algorithms due to its relatively low memory requirement [45], as well by using dropout, which is a method that evades extracting same features over and over again to reduce the risk of overfitting [46]. For the NN, FFNN, CFNN, and RNN classifiers, the number of selected hidden layers is equal to 10 and the number of neurons in the hidden layers is 50. For the CNN classifier, we used a convolution layer, ReLU function, pooling layer, fully connected layer, and softmax layer. Furthermore, to train the neural network, CNN uses the cross-entropy loss function. Moreover, Adam optimization algorithm is applied.

**Table 3.** Hyperparameter settings.

Hyperparameters	Values
Optimizer	Adam
Loss function	Cross-entropy
Dropout	0.2
Learning rate	0.001
Regularizer	L2
Maximum epochs	20
Mini-batch size	250
BiLSTM layer nodes	50

### 4.3. Fault Classification Results

For the purpose of performing the different experiments for FDD purposes, ten variable measurements are gathered, as listed in Table 4. These variables represent 1 healthy (at-

tached to class  $C_0$ ) and 15 faulty operating conditions of WECC (attached to  $C_i; i = 1, \dots, 15$ ), as shown in Table 5. Each mode behavior is adequately described over 2000 10-spaced samples, with 20 KHz as the sampling frequency for the training phase. We used of the samples 80% for the training phase and 20% for the testing phase.

**Table 4.** Labeling and description of the measured and monitored system variables.

Variables	Descriptions
$x_1$	$C_m$ : Mechanical torque ( $N_m$ )
$x_2$	$N_g$ : Generator speed ( $tr/m$ )
$x_3$	$i_{saq}$ : Generator current phase a (A)
$x_4$	$i_{sbg}$ : Generator current phase b (A)
$x_5$	$i_{scg}$ : Generator current phase c (A)
$x_6$	$V_{DC}$ : Bus voltage (V)
$x_7$	$P_{out}$ : Output power (W)
$x_8$	$i_{sar}$ : Grid current phase a (A)
$x_9$	$i_{sbr}$ : Grid current phase b (A)
$x_{10}$	$i_{sbr}$ : Grid current phase b (A)

**Table 5.** Creation of database for fault diagnosis system.

Classes	Mode	Training Data	Testing Data
$C_0$	<i>Healthy</i>	10,000	2500
$C_1$	<i>SFGS</i>	10,000	2500
$C_2$		10,000	2500
$C_3$		10,000	2500
$C_4$	<i>SFGrS</i>	10,000	2500
$C_5$		10,000	2500
$C_6$		10,000	2500
$C_7$	<i>MxF</i>	10,000	2500
$C_8$		10,000	2500
$C_9$		10,000	2500
$C_{10}$	<i>MFGS</i>	10,000	2500
$C_{11}$		10,000	2500
$C_{14}$		10,000	2500
$C_{12}$	<i>MFGGrS</i>	10,000	2500
$C_{13}$		10,000	2500
$C_{15}$		10,000	2500

In this paper, various classifiers are applied and the best classifier is selected on the basis of its classification accuracy. Table 6 illustrates the global performance accuracy.

In Scenario 1, the faults occur in the grid-side converter, which do not seriously affect the behavior of the wind system. In this case, all the developed techniques have showed high diagnosis performance except for CNN. However, in Scenario 2, different faults are presented in the generator-side converter that considerably affect the behavior of the system. This affects the diagnosis performance of the applied FDD techniques. In cases of Scenarios 3, 4, and 5, the FDD techniques showed good results, with the exception of CNN. When dealing with all fault scenarios, it is clear from Tables 6 and 7 that the BiLSTM classifier provides better classification performance when compared to the classical methods.

**Table 6.** Performance comparison of conventional techniques.

Fault Side	Techniques	Global Performance					
		Accuracy	Recall	Precision	F1 Score	CT (s)	PL
SFGS	NN	72.26	71.87	64.82	68.6	0.21	**
SFGrS		96.94	96.94	96.95	96.94	0.18	***
MFGS		87.41	88.22	88.54	88.38	0.20	***
MFGrS		97.16	97.39	97.88	97.63	0.21	****
MxF		93.79	94.4	95.42	94.71	0.18	***
All faults		59.62	58.53	59.87	59.19	0.35	*
SFGS	FFNN	76.94	76.94	78.89	77.90	0.15	**
SFGrS		87.89	87.89	89.74	88.80	0.18	***
MFGS		84.81	84.81	88.35	86.54	0.20	***
MFGrS		85.35	85.35	85.59	85.47	0.14	***
MxF		95.14	95.14	95.93	95.53	0.15	***
All faults		49.46	45.72	44.71	45.21	0.34	*
SFGS	CFNN	75.11	75.11	75.3	75.20	0.18	**
SFGrS		96.75	96.75	96.79	96.67	0.18	***
MFGS		90.46	90.46	91.32	90.88	0.14	***
MFGrS		87.78	87.78	87.81	87.79	0.24	***
MxF		95.18	95.18	95.96	95.57	0.15	***
All faults		59.87	59.56	56.75	58.12	0.18	*
SFGS	RNN	70.07	70.04	61.14	65.25	0.15	****
SFGrS		95.73	95.73	95.73	95.73	0.17	***
MFGS		80.56	80.56	75.86	78.14	0.17	***
MFGrS		86.23	86.23	87.82	87.2	0.14	***
MxF		94.23	94.47	95.47	95.47	0.17	***
All faults		47.50	47.55	40.43	43.65	0.34	*

**Table 7.** Performance comparison of deep learning techniques.

Fault Side	Techniques	Global Performance					
		Accuracy	Recall	Precision	F1 Score	CT (s)	PL
SFGS	CNN	41.16	40.33	60.43	48.37	0.95	*
SFGrS		44.06	44.06	62.01	51.52	0.74	*
MFGS		52.85	52.85	53.10	52.97	0.7	*
MFGGrS		38.18	38.18	38.74	38.46	0.77	*
MxF		42.55	42.55	44.89	43.68	0.95	*
All faults		16.43	14.39	11.98	13.07	1.29	*
SFGS	LSTM	75.08	75.08	75.21	75.14	0.71	**
SFGrS		86.90	86.9	88.96	87.92	0.55	***
MFGS		88.40	88.40	89.48	88.94	0.67	***
MFGGrS		91.83	91.83	93.18	92.50	0.59	***
MxF		100.0	100.0	100.0	100.0	0.81	****
All faults		73.70	73.58	73.70	73.64	1.32	**
SFGS	BiLSTM	72.92	65.77	72.5	72.60	0.78	**
SFGrS		95.71	95.71	95.72	88.24	0.7	***
MFGS		88.49	88.49	89.66	89.07	0.64	***
MFGGrS		90.71	90.71	93.19	91.93	0.78	***
MxF		100.0	100.0	100.0	100.0	0.75	****
All faults		79.0	79.54	81.34	81.34	1.62	**



In order to improve further the above results, a novel FDD approach is proposed using a KPCA-based BiLSTM, in which the most informative features are extracted and selected using KPCA and then fed to the BiLSTM for fault classification purposes.

As shown in Table 8, the developed KPCA-BiLSTM approach reached an accuracy rate of 97.20%. This result demonstrates its enhanced classification performance when compared to the standard BiLSTM.

**Table 8.** Performance comparison of different techniques.

Fault Side	Techniques	Global Performance					
		Accuracy	Recall	Precision	F1 Score	CT (s)	PL
All faults	<b>BiLSTM</b>	79.0	79.54	81.34	80.43	1.62	**
All faults	<b>KPCA-BiLSTM</b>	97.20	97.20	97.30	97.25	2.56	****

To better assess the efficiency of the proposed approach, the testing classification results are illustrated in Table 9 using the confusion matrix (CM). The CM illustrates the correctly classified samples and misclassified ones for the healthy case ( $C_0$ ) and faulty cases ( $C_1$  to  $C_{15}$ ). For example, for the healthy case ( $C_0$ ), the KPCA-based BiLSTM approach determined 2320 observations among 2500 (true positive). For this class, 7.2% were misclassifications (false alarms). In the faulty operating modes ( $C_5$ ,  $C_8$ ,  $C_9$ ,  $C_{11}$ ,  $C_{12}$ ,  $C_{13}$ ), the precision was 100% and the recall was 100%, with 0.0% misclassification.

**Table 9.** Confusion matrix of KPCA-based BiLSTM in testing phase.

True Classes	Predicted Classes																Recall
	$C_0$	$C_1$	$C_2$	$C_3$	$C_4$	$C_5$	$C_6$	$C_7$	$C_8$	$C_9$	$C_{10}$	$C_{11}$	$C_{12}$	$C_{13}$	$C_{14}$	$C_{15}$	
$C_0$	2320	0	0	0	110	0	0	0	0	0	0	0	0	0	0	0	92.80
$C_1$	23	2308	23	32	0	2	0	105	0	7	0	0	0	0	0	0	92.32
$C_2$	0	0	2387	0	8	22	22	5	2	0	41	13	0	0	0	0	95.48
$C_3$	9	4	0	2428	0	18	0	0	0	0	0	5	8	9	0	19	97.12
$C_4$	129	0	0	0	2309	0	0	0	0	0	0	0	0	0	0	62	92.36
$C_5$	0	0	0	0	0	2500	0	0	0	0	0	0	0	0	0	0	100.0
$C_6$	3	0	0	0	0	0	2487	0	0	0	0	0	0	0	0	10	99.48
$C_7$	213	0	0	0	14	1	0	2185	0	0	0	0	0	0	0	87	87.40
$C_8$	0	0	0	0	0	0	0	0	2500	0	0	0	0	0	0	0	100.0
$C_9$	0	0	0	0	0	0	0	0	0	2500	0	0	0	0	0	0	100.0
$C_{10}$	0	3	0	0	0	0	0	0	0	0	2497	0	0	0	0	0	99.88
$C_{11}$	0	0	0	0	0	0	0	0	0	0	0	2500	0	0	0	0	100.0
$C_{12}$	0	0	0	0	0	0	0	0	0	0	0	0	2500	0	0	0	100.0
$C_{13}$	0	0	0	0	0	0	0	0	0	0	0	0	0	2500	0	0	100.0
$C_{14}$	0	0	0	0	0	0	0	0	0	0	0	0	1	0	2499	0	99.96
$C_{15}$	29	0	0	0	11	0	0	0	0	0	0	0	0	0	0	2500	100.0
Precision	85.10	99.69	99.04	98.69	94.16	98.30	99.12	95.20	99.92	99.72	98.38	99.28	99.64	99.64	100	93.35	<b>97.30</b>

## 5. Conclusions and Future Works

In this paper, an enhanced KPCA-based BiLSTM method was presented for wind energy conversion (WEC) system fault detection and diagnosis (FDD). The proposed FDD approach was addressed in such a way that the extracted and selected features using the KPCA model are introduced as input for the BiLSTM for classification purposes. In fact, the effectiveness of the proposed classifier was validated by comparing it with several classical methods, including NN, FFNN, CFNN, RNN, and CNN. In order to evaluate the performance of the developed KPCA-based BiLSTM approach, we used data obtained from healthy WEC converters (WECC) that were then injected with several fault scenarios of fault: simple fault generator-side, simple fault grid-side, multiple faults generator-side, multiple faults grid side, and mixed faults both side. The obtained results showed the effectiveness and robustness of the proposed FDD approach in terms of accuracy, recall, precision, and computation time. The fault diagnosis accuracy when using the proposed tools showed

some missed detection and false alarm results, and some faults were not correctly classified. Thus, one future research direction is to develop adaptive BiLSTM-based tools to update the model in order to reduce missed classification results. Another future direction is to develop adaptive BiLSTM-based approaches dealing with uncertainties in WTC systems using interval-valued data representation. Additionally, ensemble-based models will be developed using multiple models in order to enhance decision-making accuracy. Ensemble-based models merges multiple learning models in order to produce one optimal predictive model that gives effective diagnosis results. Furthermore, in this study, we considered a wind profile where the mean value of the speed, as well as the pitch angle, is constant. In the real world, the wind has a variable profile according to climatic conditions. Thus, one future research direction is to implement an FDD approach while taking into account wind variations.

**Author Contributions:** Investigation, K.A.; Methodology, Z.Y., M.H. and M.M.; Supervision, M.H., M.M., K.B. and H.N.; Validation, M.H. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was supported by the Qatar National Library through the Qatar National Research Fund (QNRF) Research Grant.

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

## Abbreviations

FDD	Fault Detection and Diagnosis
FES	Feature Extraction and Selection
PCA	Principal Component Analysis
KPCA	Kernel PCA
LSTM	Long Short Term Memory
BiLSTM	Bidirectional LSTM
CNN	Convolutional Neural Network
WEC	Wind Energy Conversion
WECC	WEC Converters
WT	Wind Turbine
CPV	Cumulative Percentage of Variance
CT	Computation Time
CM	Confusion Matrix

## References

1. Kouadri, A.; Hajji, M.; Harkat, M.F.; Abodayeh, K.; Mansouri, M.; Nounou, H.; Nounou, M. Hidden Markov model based principal component analysis for intelligent fault diagnosis of wind energy converter systems. *Renew. Energy* **2020**, *150*, 598–606. [\[CrossRef\]](#)
2. Fezai, R.; Dhibi, K.; Mansouri, M.; Trabelsi, M.; Hajji, M.; Bouzrara, K.; Nounou, H.; Nounou, M. Effective Random Forest-Based Fault Detection and Diagnosis for Wind Energy Conversion Systems. *IEEE Sens. J.* **2020**, *21*, 6914–6921. [\[CrossRef\]](#)
3. Faulstich, S.; Hahn, B.; Tavner, P.J. Wind turbine downtime and its importance for offshore deployment. *Wind Energy* **2011**, *14*, 327–337. [\[CrossRef\]](#)
4. Zhang, Y.; Zheng, H.; Liu, J.; Zhao, J.; Sun, P. An anomaly identification model for wind turbine state parameters. *J. Clean. Prod.* **2018**, *195*, 1214–1227. [\[CrossRef\]](#)
5. Ravikumar, K.; Subbiah, R.; Ranganathan, N.; Bensingh, J.; Kader, A.; Nayak, S.K. A review on fatigue damages in the wind turbines: Challenges in determining and reducing fatigue failures in wind turbine blades. *Wind Eng.* **2020**, *44*, 434–451. [\[CrossRef\]](#)
6. Mishnaevsky, L. Root Causes and Mechanisms of Failure of Wind Turbine Blades: Overview. *Materials* **2022**, *15*, 2959. [\[CrossRef\]](#)
7. Singh, G.; Sundaram, K. Methods to improve wind turbine generator bearing temperature imbalance for onshore wind turbines. *Wind Eng.* **2021**, *46*, 0309524X211015292. [\[CrossRef\]](#)
8. Xu, Y.; Nascimento, N.M.M.; de Sousa, P.H.F.; Nogueira, F.G.; Torrico, B.C.; Han, T.; Jia, C.; Rebouças Filho, P.P. Multi-sensor edge computing architecture for identification of failures short-circuits in wind turbine generators. *Appl. Soft Comput.* **2021**, *101*, 107053. [\[CrossRef\]](#)
9. Xiao, C.; Liu, Z.; Zhang, T.; Zhang, X. Deep learning method for fault detection of wind turbine converter. *Appl. Sci.* **2021**, *11*, 1280. [\[CrossRef\]](#)

10. López-Uruñuela, F.J.; Fernández-Díaz, B.; Pagano, F.; López-Ortega, A.; Pinedo, B.; Bayón, R.; Aguirrebeitia, J. Broad review of “White Etching Crack” failure in wind turbine gearbox bearings: Main factors and experimental investigations. *Int. J. Fatigue* **2021**, *145*, 106091. [\[CrossRef\]](#)
11. Wang, L.; Zhang, Z.; Long, H.; Xu, J.; Liu, R. Wind turbine gearbox failure identification with deep neural networks. *IEEE Trans. Ind. Inform.* **2016**, *13*, 1360–1368. [\[CrossRef\]](#)
12. Dao, P.B.; Staszewski, W.J.; Barszcz, T.; Uhl, T. Condition monitoring and fault detection in wind turbines based on cointegration analysis of SCADA data. *Renew. Energy* **2018**, *116*, 107–122. [\[CrossRef\]](#)
13. Qiao, W.; Lu, D. A survey on wind turbine condition monitoring and fault diagnosis—Part I: Components and subsystems. *IEEE Trans. Ind. Electron.* **2015**, *62*, 6536–6545. [\[CrossRef\]](#)
14. Mansouri, M.; Fezai, R.; Trabelsi, M.; Nounou, H.; Nounou, M.; Bouzrara, K. Reduced Gaussian process regression based random forest approach for fault diagnosis of wind energy conversion systems. *IET Renew. Power Gener.* **2021**, *15*, 3612–3621. [\[CrossRef\]](#)
15. Kamel, T.; Biletskiy, Y.; Chang, L. Fault diagnoses for industrial grid-connected converters in the power distribution systems. *IEEE Trans. Ind. Electron.* **2015**, *62*, 6496–6507. [\[CrossRef\]](#)
16. Kamel, T.; Biletskiy, Y.; Chang, L. Real-time diagnosis for open-circuited and unbalance faults in electronic converters connected to residential wind systems. *IEEE Trans. Ind. Electron.* **2015**, *63*, 1781–1792. [\[CrossRef\]](#)
17. Jlassi, I.; Estima, J.O.; El Khil, S.K.; Bellaaj, N.M.; Cardoso, A.J.M. Multiple open-circuit faults diagnosis in back-to-back converters of PMSG drives for wind turbine systems. *IEEE Trans. Power Electron.* **2014**, *30*, 2689–2702. [\[CrossRef\]](#)
18. Duan, P.; Xie, K.g.; Zhang, L.; Rong, X. Open-switch fault diagnosis and system reconfiguration of doubly fed wind power converter used in a microgrid. *IEEE Trans. Power Electron.* **2010**, *26*, 816–821. [\[CrossRef\]](#)
19. Abid, A.; Khan, M.T.; Iqbal, J. A review on fault detection and diagnosis techniques: Basics and beyond. *Artif. Intell. Rev.* **2021**, *54*, 3639–3664. [\[CrossRef\]](#)
20. Zeroual, A.; Harrou, F.; Sun, Y.; Messai, N. Integrating model-based observer and Kullback–Leibler metric for estimating and detecting road traffic congestion. *IEEE Sens. J.* **2018**, *18*, 8605–8616. [\[CrossRef\]](#)
21. Nai-Zhi, G.; Ming-Ming, Z.; Bo, L. A data-driven analytical model for wind turbine wakes using machine learning method. *Energy Convers. Manag.* **2022**, *252*, 115130. [\[CrossRef\]](#)
22. Simani, S.; Farsoni, S.; Castaldi, P. Data-driven techniques for the fault diagnosis of a wind turbine benchmark. *Int. J. Appl. Math. Comput. Sci.* **2018**, *28*, 247–268. [\[CrossRef\]](#)
23. Pashazadeh, V.; Salmasi, F.R.; Araabi, B.N. Data driven sensor and actuator fault detection and isolation in wind turbine using classifier fusion. *Renew. Energy* **2018**, *116*, 99–106. [\[CrossRef\]](#)
24. Hsu, J.Y.; Wang, Y.F.; Lin, K.C.; Chen, M.Y.; Hsu, J.H.Y. Wind turbine fault diagnosis and predictive maintenance through statistical process control and machine learning. *IEEE Access* **2020**, *8*, 23427–23439. [\[CrossRef\]](#)
25. Wu, Y.; Fu, Z.; Fei, J. Fault diagnosis for industrial robots based on a combined approach of manifold learning, treelet transform and Naive Bayes. *Rev. Sci. Instrum.* **2020**, *91*, 015116. [\[CrossRef\]](#)
26. Tuerxun, W.; Chang, X.; Hongyu, G.; Zhijie, J.; Huajian, Z. Fault diagnosis of wind turbines based on a support vector machine optimized by the sparrow search algorithm. *IEEE Access* **2021**, *9*, 69307–69315. [\[CrossRef\]](#)
27. Lu, J.; Qian, W.; Li, S.; Cui, R. Enhanced K-nearest neighbor for intelligent fault diagnosis of rotating machinery. *Appl. Sci.* **2021**, *11*, 919. [\[CrossRef\]](#)
28. Jalayer, M.; Orsenigo, C.; Vercellis, C. Fault detection and diagnosis for rotating machinery: A model based on convolutional LSTM, Fast Fourier and continuous wavelet transforms. *Comput. Ind.* **2021**, *125*, 103378. [\[CrossRef\]](#)
29. He, Z.; Shao, H.; Zhong, X.; Zhao, X. Ensemble transfer CNNs driven by multi-channel signals for fault diagnosis of rotating machinery cross working conditions. *Knowl.-Based Syst.* **2020**, *207*, 106396. [\[CrossRef\]](#)
30. Xiang, L.; Wang, P.; Yang, X.; Hu, A.; Su, H. Fault detection of wind turbine based on SCADA data analysis using CNN and LSTM with attention mechanism. *Measurement* **2021**, *175*, 109094. [\[CrossRef\]](#)
31. Han, T.; Li, Y.F.; Qian, M. A hybrid generalization network for intelligent fault diagnosis of rotating machinery under unseen working conditions. *IEEE Trans. Instrum. Meas.* **2021**, *70*, 1–11. [\[CrossRef\]](#)
32. Mansouri, M.; Dhibi, K.; Hajji, M.; Bouzara, K.; Nounou, H.; Nounou, M. Interval-Valued Reduced RNN for Fault Detection and Diagnosis for Wind Energy Conversion Systems. *IEEE Sens. J.* **2022**. [\[CrossRef\]](#)
33. Mansouri, M.; Trabelsi, M.; Nounou, H.; Nounou, M. Deep learning based fault diagnosis of photovoltaic systems: A comprehensive review and enhancement prospects. *IEEE Access* **2021**, *9*, 126286–126306. [\[CrossRef\]](#)
34. Lei, J.; Liu, C.; Jiang, D. Fault diagnosis of wind turbine based on Long Short-term memory networks. *Renew. Energy* **2019**, *133*, 422–432. [\[CrossRef\]](#)
35. Lee, J.M.; Yoo, C.; Choi, S.W.; Vanrolleghem, P.A.; Lee, I.B. Nonlinear process monitoring using kernel principal component analysis. *Chem. Eng. Sci.* **2004**, *59*, 223–234. [\[CrossRef\]](#)
36. Zhang, Y. Enhanced statistical analysis of nonlinear processes using KPCA, KICA and SVM. *Chem. Eng. Sci.* **2009**, *64*, 801–811. [\[CrossRef\]](#)
37. Thara, D.; PremaSudha, B.; Xiong, F. Epileptic seizure detection and prediction using stacked bidirectional long short term memory. *Pattern Recognit. Lett.* **2019**, *128*, 529–535.
38. Zhao, R.; Yan, R.; Wang, J.; Mao, K. Learning to monitor machine health with convolutional bi-directional LSTM networks. *Sensors* **2017**, *17*, 273. [\[CrossRef\]](#)

39. Faust, O.; Shenfield, A.; Kareem, M.; San, T.R.; Fujita, H.; Acharya, U.R. Automated detection of atrial fibrillation using long short-term memory network with RR interval signals. *Comput. Biol. Med.* **2018**, *102*, 327–335. [[CrossRef](#)]
40. Hoffmann, H. Kernel PCA for novelty detection. *Pattern Recognit.* **2007**, *40*, 863–874. [[CrossRef](#)]
41. Dhibi, K.; Mansouri, M.; Bouzrara, K.; Nounou, H.; Nounou, M. An Enhanced Ensemble Learning-Based Fault Detection and Diagnosis for Grid-Connected PV Systems. *IEEE Access* **2021**, *9*, 155622–155633. [[CrossRef](#)]
42. Valle, S.; Li, W.; Qin, S.J. Selection of the number of principal components: The variance of the reconstruction error criterion with a comparison to other methods. *Ind. Eng. Chem. Res.* **1999**, *38*, 4389–4401. [[CrossRef](#)]
43. Yang, Z.; Chai, Y. A survey of fault diagnosis for onshore grid-connected converter in wind energy conversion systems. *Renew. Sustain. Energy Rev.* **2016**, *66*, 345–359. [[CrossRef](#)]
44. Hameed, Z.; Garcia-Zapirain, B. Sentiment classification using a single-layered BiLSTM model. *IEEE Access* **2020**, *8*, 73992–74001. [[CrossRef](#)]
45. Aslan, M.F.; Unlarsen, M.F.; Sabanci, K.; Durdu, A. CNN-based transfer learning–BiLSTM network: A novel approach for COVID-19 infection detection. *Appl. Soft Comput.* **2021**, *98*, 106912. [[CrossRef](#)]
46. Xia, M.; Li, T.; Xu, L.; Liu, L.; De Silva, C.W. Fault diagnosis for rotating machinery using multiple sensors and convolutional neural networks. *IEEE/ASME Trans. Mechatron.* **2017**, *23*, 101–110. [[CrossRef](#)]