Fully Automated Diagnostics of Induction Motor Drives in Offshore Wind Turbine Pitch Systems using Extended Park Vector Transform and Convolutional Neural Network

Manuel S Mathew[1], Surya Teja Kandukuri[2], Christian W Omlin[3]

[1,3]*University of Agder, Jon Lilletuns vei 9, 4879 Grimstad, Norway*
manuel.s.mathew@uia.no
christian.omlin@uia.no

[2]*Norwegian Research Centre, Energy & Technology Department, Tullins Gate 2, 0166 Oslo, Norway*
suka@norceresearch.no

ABSTRACT

Due to their location and related complexities, the offshore wind farms (OWF) have higher downtimes and operation and maintenance (O&M) costs compared to their onshore counterparts. Condition monitoring could help in bringing down the O&M costs of OWFs. The pitch system is one of the components most prone to failure. This paper details an approach for enhanced diagnosis of the electric pitch systems especially focusing on the induction motor drives (IMD) in wind turbines. The proposed method uses an extended Park vector approach (EPVA) in conjunction with a convolutional neural network (CNN) to accurately classify the condition of an IMD and localize the faults. The method is validated on data collected from a laboratory setup. The advantage of the proposed approach is that the condition of the IMD can accurately be classified, and faults localized in operating conditions with varying load and frequency without any additional information on the instantaneous operating speed, frequency, or load on the motor drives. This results in a non-invasive diagnostic approach incurring least additional expenses to implement.

1. INTRODUCTION

Offshore wind farms (OWF) have significant potential to contribute towards global energy sustainability. However, they face unique operational challenges, mainly because of their remote locations and harsh marine environments in which they operate. To put this in perspective, while onshore wind farms are attaining a 95% to 97% availability for modern systems Pfaffel, Faulstich and Rohrig (2017), the availability of OWFs is relatively lower and highly variable. The data from earlier offshore wind farms suggest an availability of 67% to 85% (Feng, Tavner, & Long, 2010) with more latest estimates of 80% to 84% (Cevasco, Koukoura, & Kolios, 2021). The limited weather windows for performing necessary maintenance leads to longer downtimes, which explains the gap in operational availability between onshore and offshore wind farms. Furthermore, the operational and maintenance (O&M) costs constitute a significant proportion of the lifetime costs associated with OWFs with estimates ranging from roughly 23% on the lower end (Ren, Verma, Li, Teuwen, & Jiang, 2021) to 30% at the higher end of the spectrum (Hammond, & Cooperman, 2022). This represents a stark contrast to onshore wind farms, where the lifetime O&M costs typically account for approximately 5% (Ren et al., 2021). Thus, implementing condition-based maintenance (CBM) strategies, and hence condition monitoring (CM) become vital in reducing the costs associated with O&M and helps in reducing the downtimes in maintenance activities.

The pitch system of wind turbines is among the components most prone to failures and one that contributes significantly to a non-trivial amount of downtime. The results from ReliaWind project (Wilkinson et al., 2010) indicate the pitch system was responsible for nearly 15% of failures per turbine per year and close to 20% of total downtime hours per year across different manufacturers in their database. A more recent study (Walgern, Fischer, Hentschel, & Kolios, 2023) suggests a pitch system failure rate of 0.54 (hydraulic) and 0.56 (electrical) per turbine per year. Pitch systems are also found to be the most critical subcomponent in the premature failure period (Santelo, De Oliveira, Maciel, & De A. Monteiro, 2022). This makes the pitch system an ideal candidate for enabling CM systems, particularly in OWFs because of the additional costs and downtimes associated with their reactive maintenance.

Although there have been some efforts to develop CM solutions for wind turbine pitch systems, the extent of these attempts has not been commensurate with their critical impact on downtime and failure rates. Cho, Gao and Moan (2016) developed a Kalman filter based method for diagnosing pitch sensor and actuator faults in floating wind turbines based on the NREL 5MW wind turbine model. However, the focus here was not on the incipient fault detection. Several machine learning-based techniques for fault diagnostics in pitch systems are also found in the literature. supervisory control and data acquisition (SCADA) data has been used to detect anomalies in the pitch system in tandem with Isolation Forest based anomaly detection models (Mckinnon, Carroll, Mcdonald, Koukoura, & Plumley, 2021). In this work, the authors developed different models of varying amounts of training data to detect anomalous patterns. Further, they experimented with varying lengths of post-processing window to see how it affects their model. Their results show that their method could notice signs of turbine failure 12 to 18 months ahead. Park, Kim, Dinh and Park (2022) used neural networks to find abnormal operations in the pitch system of a wind turbine. The authors define the abnormal operation for the pitch system using the deviation in the blade pitch angle, where a deviation of more than 4.95° in blade pitch angle was considered abnormal. Wei, Qian and Zareipour (2019) developed a condition monitoring and fault detection system for the wind turbine pitch system using optimized relevance vector machine regression. Their work leverages the SCADA data to detect faults in the pitch system particularly focusing on encoder failures, pitch controller failures, electric motor failures, and slip ring failures. Similarly Chen, Matthews and Tavner (2013) used SCADA data to develop an a-priori knowledge-based ANFIS (APK-ANFIS) model to detect faults in the wind turbine pitch system. The authors identified four critical characteristics features (CF) of pitch faults after analyzing data in the developmental stage of a fault and that immediately after the maintenance has been carried out. These CFs have been used to develop the corresponding APK-ANFIS models, the results from which where aggregated to detect the fault in the pitch systems. Most of the studies in the literature reviewed are effective in detecting faults in the pitch systems, however, they fall short of delineating the fault diagnosis to a subcomponent level.

A subcomponent level diagnosis of faults is essential as it contributes to efficient planning and implementation of maintenance activities, especially in OWFs, where precise planning is paramount. While SCADA data can be effectively used for preliminary fault diagnosis, it is less effective in the subcomponent level fault diagnosis. In this paper, we focus on fault diagnostics for the induction motor drive (IMD) of an electrical pitch system. Subcomponent level fault diagnosis for pitch motor drives using current signature analysis have been previously addressed by the authors (Kandukuri, Karimi, & Robbersmyr, 2016; Kandukuri, Senanayaka, Huynh, Karimi, & Robbersmyr, 2017) , and also proposed a two-stage fault classification scheme based on support vector machine (SVM), for large-scale deployment in OWFs (Kandukuri, Senanyaka, & Robbersmyr, 2019).

The issue with classical current signature-based methods in fault detection is that there is an assumption of steady state operations in terms of speed and load. The wind turbine pitch systems on the other hand are operated intermittently and are exposed to varying speed and load profiles. This means that, either regions of steady state operations must be carefully detected for data acquisition or advanced signal processing techniques are to be employed (Benbouzid, M El Hachemi, 2000; Bhole, & Ghodke, 2021; Liu, & Bazzi, 2017).

Thus, in this paper, a novel solution is proposed by calculating the extended Park vector (EPV) current from the three-phase motor line currents and then extracting the time-frequency representation using Short-Term Fourier Transform (STFT). The three-phase motor line currents for this purpose are observed at varying operating conditions: speed, and load. For detecting the condition of the IMD, these representations are subsequently converted into spectrograms, which are then used to train a convolutional neural network (CNN) for classification of the IMD's condition. CNNs have earlier been reported focusing on diagnostics of gearboxes (Amin, Bibo, Panyam, & Tallapragada, 2023; Gecgel, Ekwaro-Osire, Gulbulak, & Morais, 2021; Jiang, Han, & Xu, 2020), bearings (Choudhary, Mian, & Fatima, 2021; Lu et al., 2023; Ruan, Wang, Yan, & Gühmann, 2023; Wang, Mao, & Li, 2021; Yuan, Lian, Kang, Chen, & Zhai, 2020), and IMDs (Junior et al., 2022; Khanjani, & Ezoji, 2021; Kumar, & Hati, 2022; Lee, Pack, & Lee, 2019; Skowron, Orlowska-Kowalska, Wolkiewicz, & Kowalski, 2020). However, most of these works depend on vibration sensors, or are specific to one type of fault. Further, most of them assume constant supply frequency and load. Skowron et al. (2020) warrants a special mention as they use motor line currents to detect faults in an IMD. They normalize these currents to create a vector that is then reshaped to form an RGB matrix corresponding to each of the three phases. While they were able to detect and differentiate between various kinds of stator faults including insipient faults, their work still deals with one kind of fault within the induction motor.

Thus, what differentiates this work from other works are as follows:

1. Fault diagnostics of IMDs operating under varying speed, frequency, and load conditions without needing any additional data on these parameters.
2. Beyond identifying a single type of fault, the proposed approach is capable of fault localization using extended Park vector approach (EPVA) in conjunction with a CNN classifier.

3. The proposed approach, through EPVA, negates the need for additional sensors to be deployed. This makes it an economically viable option for wind turbines without vibration sensors, especially those that are nearing their end of designed life. Because while vibration sensor-based diagnostics are more widely applicable, they are expensive (Trajin, Regnier, & Faucher, 2010) compared to motor current signature analysis (MCSA).
4. Since continuous monitoring of the WT pitch system is not required in this method, intermittent snapshots of the three-phase currents are sufficient for reliable diagnosis. Thus, reducing the data transmission load from each turbine.
5. The proposed algorithm need not be implemented at each turbine, the proposed approach can contribute towards farm-level health management.

While EPVA and similar MCSA methods have earlier been used for IM diagnosis and prognosis (Erik Leandro, Levy Ely De Lacerda De, Jonas Guedes Borges Da, Germano, & Luiz Eduardo Borges Da, 2012), to the best of authors' knowledge this is the first time the diagnostics of the induction motor has been fully automated while using spectrogram of the EPVA in conjunction with a deep learning based classifier. The rest of the paper is organized as follows. In section 2, the induction motor faults under consideration and the reason for selecting these are discussed. This is followed by a brief explanation of theories of EPVA and CNNs in section 3. Section 4 details the laboratory setup used to collect the necessary data and Section 5 discusses the methodology of research and details about the CNN based classifier. Results from the classification scheme are discussed in section 6. The paper is concluded, and possible future directions are briefly highlighted in section 7.

2. IMD FAULTS CONSIDERED

Despite being robust, induction motors are not immune to failures. Stator faults and bearing faults are among the most reported components contributing to the total failures in an IMD (Benbouzid, M., 1999; Benbouzid, M. E. H., & Kliman, 2003; Singh, & Ahmed Saleh Al Kazzaz, 2003; Thorsen, & Dalva, 1995) as shown in Figure 1. Nearly half of the total failures are the result of stator faults, making it one of the most important types of faults to be detected. This is followed by bearing faults which account for almost one-third of the total failures. Compared to those, a menial 10% of the failures are accounted for by rotor faults. These faults occur generally because of drive-generated harmonics, poor ventilation at low-speed operation, and abrupt load variations.

Thus, in this study, we have considered the two components contributing the most to the total IMD failures: stator fault and bearing fault. A future study may be done including rotor bar faults such as different severities of broken rotor bar (BRB) faults.

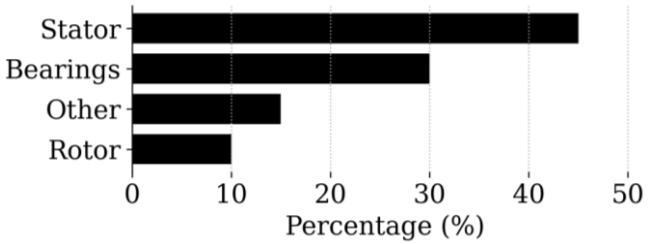

Figure 1. Distribution of IMD faults

3. THEORY

3.1. Extended Park Vector (EPV) Analysis

The theory behind using MCSA for IMDs revolves around the concept that an induction motor, while operating in healthy state, is symmetrical across the three phases. A fault in the motor disrupts this symmetry causing a periodically recurring asymmetry in the motor's operational characteristics. This periodic recurrence manifests as a particular frequency in the current known as "fault frequencies" or "signature frequencies."

These fault frequencies arise due to the interaction between the motor's electrical and mechanical components influenced by the fault. For example, a stator fault, such as insulation failure, due to short circuits between the stator windings, or phase imbalance, causes an asymmetric distribution of electromagnetic fields (EMF) within the motor. This asymmetry causes variations in the magnetic effect on the rotor resulting in irregular rotor motion. The interaction between the EMF and the rotor's motion produces specific frequency components in the motor's currents called "stator fault frequencies". These stator fault frequencies are then reflected in the MCSA as harmonics of the fundamental frequency, or appearance of specific sidebands around the fundamental frequency and its sidebands. In the case of a bearing fault, which can occur because of physical damage, wear and tear, inadequate lubrication, or external factors, leads to mechanical vibrations which modulate the EMF within the motor affecting the air gap flux density. This introduces specific frequencies in the motor's current signature called "bearing fault frequencies". The specific frequency of the bearing fault depends on several factors like bearing design, motor speed, and the nature of the fault. These signature frequencies are then used to diagnose the faults in MCSA.

EPV analysis builds upon the foundational principles of MCSA and has been used for a while now in diagnosing motor electrical faults (Cardoso, Cruz, & Fonseca, 1997). The direct (i_d) and quadrature (i_q) axis currents are initially calculated as a function of three phase motor currents (i_a, i_b, i_c) as follows:

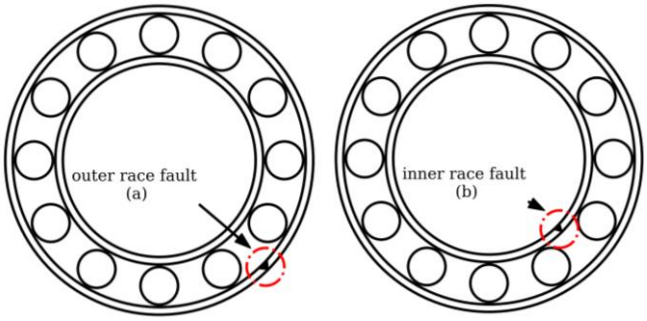

Figure 2. Schematic diagram of (a) an outer race fault, and (b) an inner race fault in a rolling element bearing.

$$i_d = \left(\sqrt{2}/\sqrt{3}\right) i_a - \left(1/\sqrt{6}\right) i_b - \left(1/\sqrt{6}\right) i_c \quad (1)$$

$$i_q = \left(1/\sqrt{2}\right) i_b - \left(1/\sqrt{2}\right) i_c \quad (2)$$

The extended Park vector (i_p) is then calculated as follows:

$$i_p = \left|\left(i_d + j i_q\right)\right| \quad (3)$$

Where, j is the imaginary unit defined as $j^2 = -1$.

When a stator turn fault (STF) occurs due to shorting between the phase windings, the three phase currents become imbalanced and the ideal values for direct and quadrature axis currents mentioned in Cardoso et al. (1997) doesn't hold anymore. The result is that the stator turn fault can be identified only using the spectral component at twice the supply frequency, f_s (Sahraoui, Zouzou, Ghoggal, & Guedidi, 2010) in the spectrum of i_p:

$$f_{STF} = 2f_s \quad (4)$$

Figure 2 shows the schematics of an outer race fault (a), and an inner race fault (b) in a rolling element bearing. A bearing fault, as discussed earlier, causes spectral components at different frequencies as determined by bearing design, motor speed, and the nature of the fault (e.g., faults on the inner race, outer race, ball spin, or cage defects). This results in three additional spectral components in the spectrum of i_p along with the fundamental component of the power supply (Zarei, & Poshtan, 2009) as:

$$f_{BRG} \in \{f_v, 2f_v, |2f_v - f_s|\} \quad (5)$$

For an outer race fault, as shown in Figure 2 (a), the characteristic vibration frequency, f_v, can be estimated using the following equation (Zarei, & Poshtan, 2009):

$$f_v \approx 0.4 N_b f_r \quad (6)$$

where N_b is the number of rolling elements in the bearing, and f_r is the shaft rotational frequency.

Even though these frequencies (f_{STF}, *and* f_{BRG}) can reliably be used for fault diagnosis in short time windows of constant operation, this fails in the case of a variable load and frequency because of the changes in the shaft rotational frequency, f_r, and supply frequency, f_s.

Characterizing the time-frequency response of the extended Park vector, i_p is critical in EPVA. The short-term Fourier transformations (STFT) is used to decompose the Park vector into its time-frequency components, which offers an in-depth view of how these frequency components evolve over time. This level of detail is more suitable for diagnosing the faults within motors operating under non-stationary conditions.

STFT is a special case of Fourier transforms where the Fourier transform is applied in series to smaller slices of the signal. The assumption here is that for a shorter time window, the original non-stationary signal becomes stationary. The STFT of a non-stationary signal *y(t)* can be estimated by discretizing the continuous-time signal to a discrete-time signal, *y(n)*, where *n* is the discrete time indices. Then the discrete STFT is calculated for the discrete-time signal as:

$$Y(\omega, b) = \sum y(n) w(n - b) e^{-j\omega n} \quad (7)$$

where $w(.)$ is the windowing function and b is the window-shifting time constant. The calculation of STFT is done using a fixed-size window, which means that if the window is longer, frequency resolution is better at the expense of time resolution and vice versa for shorter windows. Thus, deciding a window length for the STFT operation is crucial in accurately extracting the desired time-frequency information (Oppenheim, 1999) and thereby diagnosing the motor condition.

3.2. Convolutional Neural Networks (CNN)

Though there has been some precedents in computer vision research inspired by natural vision, CNNs developed by Lecun et al. (1989) were instrumental in the development of computer vision at scale. CNNs are similar to artificial neural networks or "vanilla neural networks" in that they are made up of neurons. However, CNNs generally consists of three types of layers, namely convolutional layer, pooling layer, fully connected layers, and an output layer.

The convolutional layers are used to learn a feature representation from the inputs provided to create feature maps. In this layer, a learned kernel convolves with the input producing a feature map, the result of which is then passed on to an elementwise non-linear activation function to get the activation maps. Each element of the feature map is connected to a local subset of neurons in the previous layer or the input. The feature map element at (i, j) in the k^{th} feature map of the l^{th} layer can be calculated as:

$$z^l_{i,j,k} = \boldsymbol{w}^l_k \boldsymbol{x}^l_{i,j} + b^l_k \quad (8)$$

where $\boldsymbol{w}^l_k$ and b^l_k are the weight vector and bias term of the k^{th} filter of the l^{th} layer, respectively. $\boldsymbol{x}^l_{i,j}$ is the local subset of

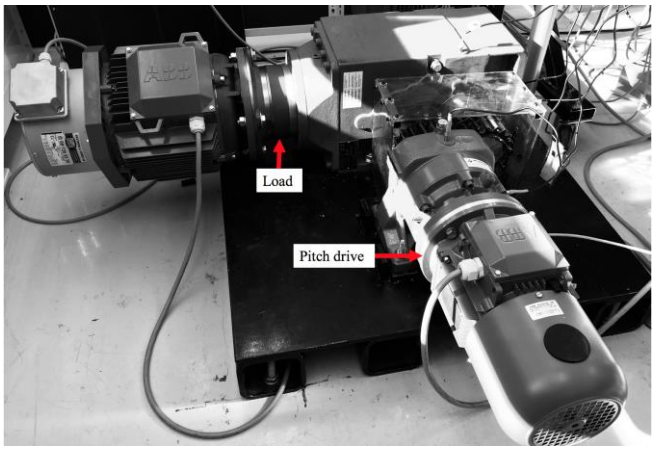

Figure 3. Laboratory setup for motor diagnostics

input to the convolutional layer centered at (i, j). An activation function such as rectified linear unit (ReLU) later introduces non-linearities helping the network to learn non-linear features.

The pooling layer, often placed between two convolutional layers, introduces shift-invariance to the feature maps. This is achieved by reducing the resolution of the feature maps. Usually, average pooling and max pooling layers are used depending on the task at hand. Higher-level feature representations are extracted eventually by stacking several convolutional and pooling layers.

One or more fully-connected layers usually succeed in CNNs aiming to achieve high level reasoning (Simonyan, & Zisserman, 2014). The last layer of CNNs is an output layer, which uses a task appropriate activation function such as sigmoid function for classification or ReLU for regression problems.

4. LABORATORY SETUP

Figure 3 shows the laboratory setup built to study the common faults in the pitch motor drives and planetary gear boxes of a wind turbine. A 1.1 kW, three-phase induction motor served as the test motor. Another 2.2 kW three-phase induction motor was used to supply the loads in the setup through a bevel-planetary-helical gearbox. Both the motors were driven by commercial field-oriented control (FOC) drives. Further details of the setup can be found in Table 1. The selection of the current sensor for this setup has been influenced by the desire for a common industrial sensor which is economical for installation on multiple units. Further, the overall frequency content of the signal has been tested over ideal power source and showed excellent signal-to-noise ratio. The speed and torque references for both test and load motors are provided to their respective FOC drives through a PC.

Table 1. Details of the test setup

Test Motor	
IM Rated Power	1.1 kW
IM Rated Speed	1420 rpm
IM Rated Torque	7.2 Nm
Current Sensor	
Model	LEM LTS-6NP
Primary nominal RMS current, I_{PN}	6 A
Accuracy @ I_{PN}, 25° C	±0.2
Data Acquisition	NI USB DAQ
Acquisition rate	15 kHz

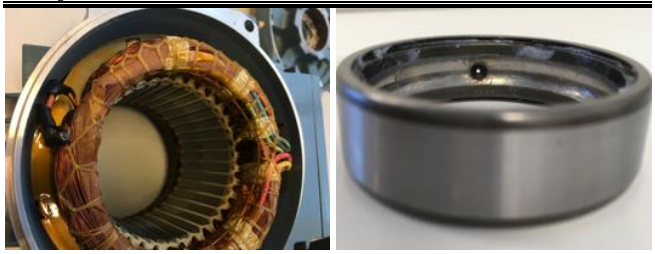

Figure 4. Seeded motor faults: stator turns fault (left), bearing fault (right).

The STF, and BRG faults were artificially seeded as shown in Figure 4. The STF was seeded by shorting 10% of a phase winding, while BRG was seeded as an outer race fault with a diameter of a ≈ 2mm through hole.

5. METHODOLOGY

Initially, the test motor was run in healthy condition, across a range of speeds varying between 850 rpm and 1420 rpm. At the same time, the loads were varied between no-load and full-load conditions at random to simulate the random loading on the wind turbine's pitch system. The speed interval was used after consulting the motor loadability curves to ensure that the motor reaches neither an overload condition nor stall condition because of the random loading. The randomness of operational conditions was ensured using different random number generators with seeds refreshed every thirty second interval. The three-phase currents were recorded as snapshots of 30 seconds each. Similarly, records were made for the motor operating in faulty conditions as mentioned in Section 4. A total of nearly a thousand minutes of data was collected from the test setup for this purpose.

Further, the Park vector, i_p and its STFT has been calculated for each of the recorded snapshots using equations (1), (2), (3), and (7). Examples of the STFT results from each of the three conditions: healthy, STF, and BRG is shown in Figure 5–7. Classical signal processing methods to detect faults from the STFT of the Park vector may fail here, however, from the figure, it can be noted that there is an increase in frequency content around 100 Hz in the STF condition, which is around $2f_s$ (Figure 6). A similar increase in frequency content can

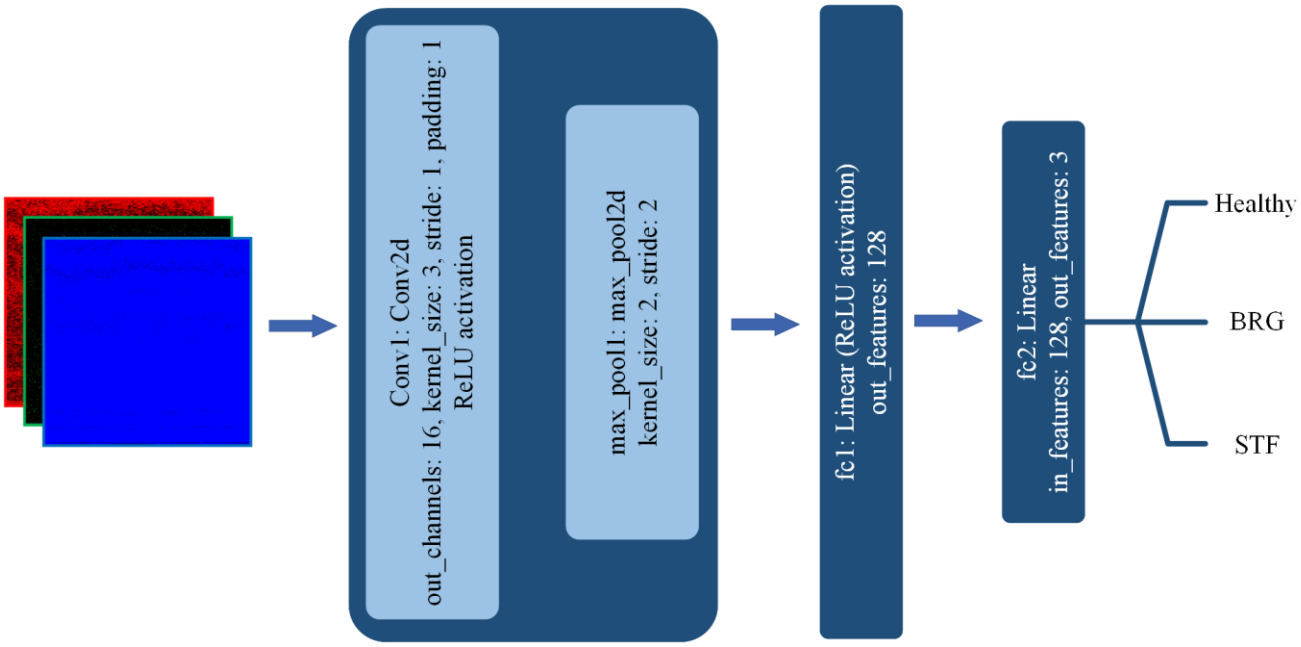

Figure 8. CNN architecture

also be observed in the case of BRG faults at around 500 – 600 Hz (Figure 7), which is distinct from the healthy case (Figure 5).

Around 2100 STFT images containing the time-frequency information associated with the operation of the motor in the three different states mentioned in the previous paragraph were then used to develop a CNN model. The entire dataset of images was split at random and 70% of the data was used in training while 15% each was used for validation and testing purposes. The STFT spectrum of the current, i_p, is obtained as an RGB image, which was then resized to 360 x 360. The architecture of the CNN that was developed for fault classification is shown in the Figure 8. The architecture consists of a convolutional layer followed by a max pooling layer, the output from which is flattened and forwarded to a fully connected layer. This fully connected layer learns high-level features from the flattened inputs. The final layer serves as the classifier, which takes the output from the fully connected layer to classify the image into one of the three conditions previously mentioned.

The CNN was trained on a system equipped with an Intel Xeon processor, NVIDIA Tesla V100-SXM3-32GB GPU on Python 3.10 using PyTorch 2.2. The model was trained over hundred epochs with a batch size of 256 and a learning rate of 10^{-3}. The learning rate was arrived at after narrowing down the value by using a learning decay scheduler. Early stopping and L2 regularization were employed to mitigate the possibility of model overfitting to the training dataset. The early stopping algorithm checks for any improvements in the validation loss and stops training if no improvement is

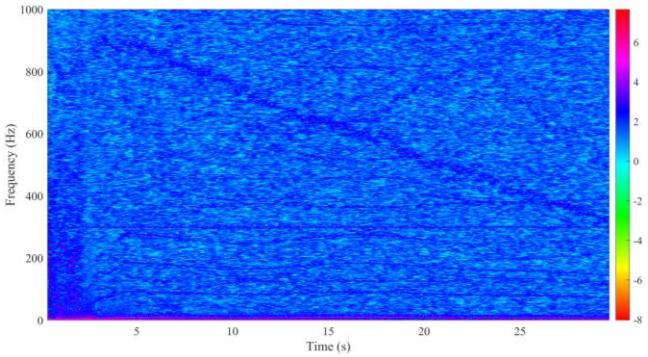

Figure 5. STFT of a Park vector, i_p, of motor working in healthy condition.

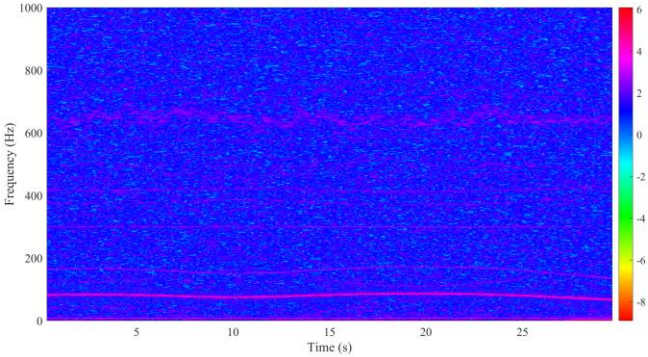

Figure 6. STFT of a Park vector, i_p, of motor with STF fault.

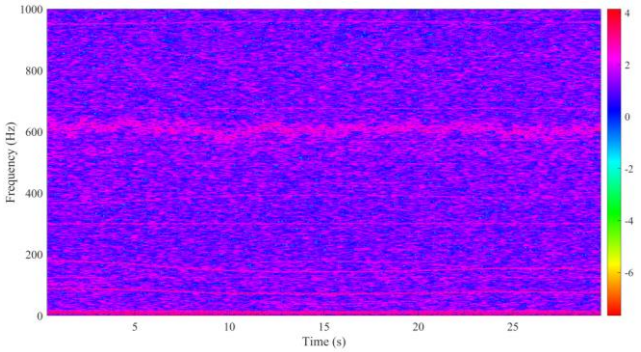

Figure 7. STFT of a Park vector, i_p, of motor with BRG fault.

observed for 25 epochs. Since this is a multi-class classification problem, the cross entropy loss was used as the loss function and the Adam optimizer (Kingma, & Ba, 2014) was selected for optimizing the loss function. At the end of each epoch the model was validated with the validation dataset and results recorded, which is detailed in the following section.

6. Results and Discussions

Figure 9 and Figure 10 illustrates the feature maps generated after the convolutional layer and max pooling layers of the trained CNN model in BRG and STF fault conditions, respectively. It is clear from the figures that the convolutional layer followed by the max pooling layer effectively identifies the specific locations within the spectrum associated with each fault condition.

After training, the model was tested on a previously unseen test dataset. Inference on GPU takes slightly higher than 17 seconds and that on CPU takes approximately 36 seconds to classify the 320 snapshots. Table 2 shows the confusion matrix of the model's performance on this dataset. The STF fault was the most accurately classified among the three

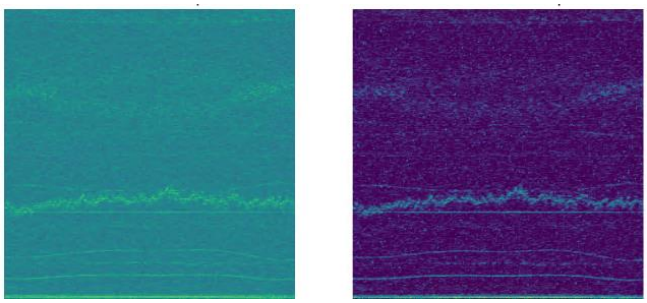

Figure 9. One of the feature maps after convolutional layer (left) and max pool layer (right) in BRG fault condition.

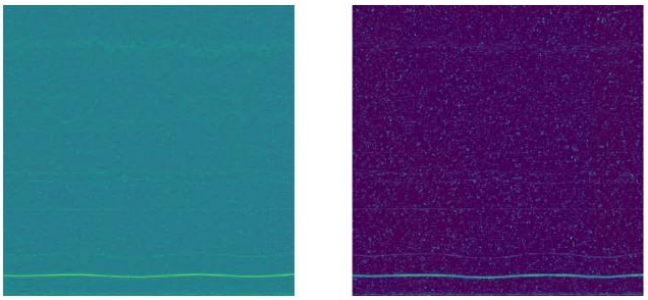

Figure 10. One of the feature maps after convolutional layer (left) and max pool layer (right) in STF fault condition.

Table 2. Confusion matrix of the model performance on test data.

	Healthy	**STF**	**BRG**	**Percentage**
Healthy	94	0	3	96.9%
STF	0	109	0	100%
BRG	4	0	110	96.5%

conditions with 100% of the cases being correctly identified as such. On the other hand, the model makes some mistakes while classifying the healthy and BRG fault conditions.

Table 3 shows the performance of the developed model in classifying the three motor conditions. The model has an overall accuracy of 97.8%. Similarly high values of precision, recall, and F1-score are observed when the model encounters the test dataset. Such high numbers might raise the suspicion of overfitting or data leakage, which is the case where the test dataset was inadvertently used in training the model. Early stopping and L2 Regularization help in preventing overfitting of the model on the training dataset while the entire data pipeline has been verified manually to ensure that there is no data leakage. Thus, the performance as shown in Table 3 highlights that the model has effectively learned the underlying pattern and can classify the three motor conditions effectively. Further, these high values in performance also indicate that the model is sufficiently complex to match the problem's complexity.

Table 3. Performance of the CNN model on test dataset.

Metric	**Healthy**	**STF**	**BRG**	**Overall**
Accuracy	-	-	-	97.8%
Precision	95.9%	100%	97.3%	97.7%
Recall	96.9%	100%	96.5%	97.8%
F1-Score	96.4%	100%	96.9%	97.78%

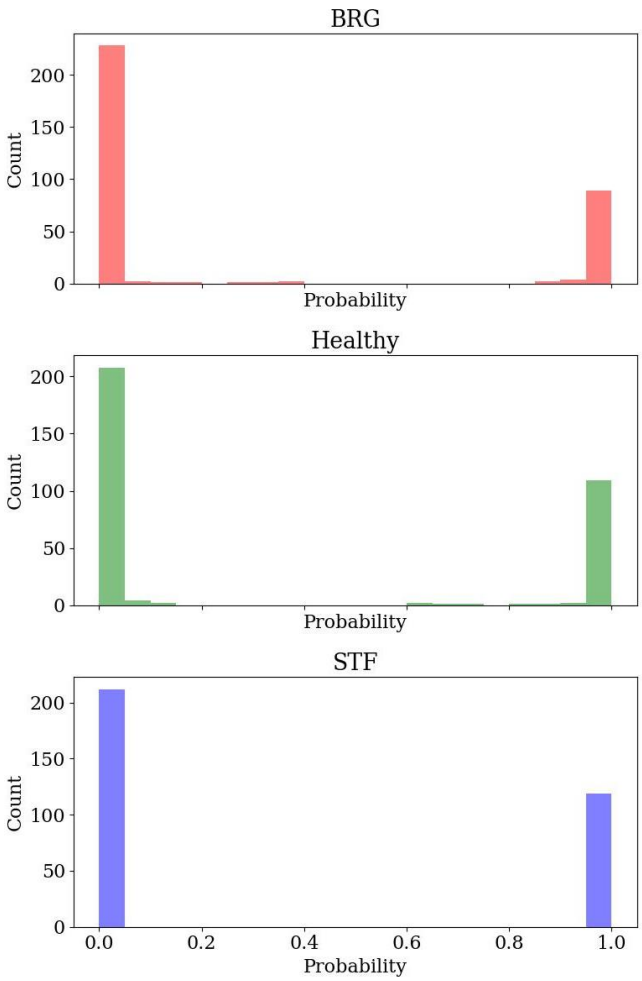

Figure 11. Histogram of predicted probabilities for each motor condition in test dataset.

To further explore the inference pattern, a histogram of the predicted probabilities was drawn across different classes as shown in Figure 11. Histogram of predicted probabilities illustrates the confidence of the model in classification to different classes. The pronounced skewness towards the extremities of the probabilities indicates a high level of certainty in classification. This is especially true for the STF fault where the model has a hundred percent confidence if a given snapshot is indicative of an STF fault (probability = 1.0) or not (probability = 0.0). On the other hand, the model is slightly less confident in identifying the healthy condition or the bearing fault as evidenced by relatively higher variance in the model's predicted probabilities.

7. Conclusion

In this paper, we have successfully developed and demonstrated a method for enhanced diagnostics for IMDs used in wind turbine pitch system. In addition to detecting the operational state of the motor (healthy/faulty), this approach also helps in localizing the fault to the stator or the bearing of the IMD. The presented approach uses the extended Park vector approach (EPVA) and short-term Fourier transforms (STFT) to extract the time-frequency information from the three-phase induction motor currents taken as 30-second snapshots. These snapshots are then classified using a convolutional neural network. The high accuracy in classification of the conditions indicates that the model can accurately diagnose the state of the IMD in question.

The advantage of the proposed method is that CNN requires only snapshots of 30 seconds each to determine the state of operation. Other than the sampling frequency, no additional information is required about the loading conditions or the frequency of operation, making it a suitable candidate for farm-level implementation. Further, as continuous monitoring is not required in this approach, it is ideal for WT pitch systems that operate intermittently. Since only 30-second snapshots of motor currents are the requirement, the data transferred from the WTs will be minimal.

The results presented here are tested on different motors of the same type, further validation of the methodology over different test motors could help strengthen the study in future. Additionally, at present two different fault conditions have been studied, more faults like broken rotor bar (BRB) faults or different stages of the STF could also be included in future works.

While the CNN classifier can be used to determine the motor condition accurately, the bottleneck in this methodology is the STFT calculations, which are computationally intensive. Thus, as a next step, the authors intend to test different ML paradigms to bypass the STFT calculations and directly detect these conditions from the Park vector current, i_p.

Acknowledgements

This research work has been funded by Analytics for asset Integrity Management of Windfarms (AIMWind), under grant no. 312486, from Research Council of Norway (RCN).

AIMWind is collaborative research from University of Agder, Norwegian Research Center (NORCE), and TU Delft, with DNV and Origo Solutions as advisory partners.

References

Amin, A., Bibo, A., Panyam, M., & Tallapragada, P. (2023). Vibration based fault diagnostics in a wind turbine planetary gearbox using machine learning. *Wind Engineering*, *47*(1), 175-189. https://doi.org/10.1177/0309524x221123968

Benbouzid, M. (1999). Bibliography on induction motors faults detection and diagnosis. *IEEE Transactions on Energy Conversion*, *14*(4), 1065-1074.

Benbouzid, M. E. H. (2000). A review of induction motors signature analysis as a medium for faults detection.

IEEE Transactions on Industrial Electronics, *47*(5), 984-993.

Benbouzid, M. E. H., & Kliman, G. B. (2003). What stator current processing-based technique to use for induction motor rotor faults diagnosis? *IEEE Transactions on Energy Conversion*, *18*(2), 238-244. https://doi.org/10.1109/TEC.2003.811741

Bhole, N., & Ghodke, S. (2021). Motor current signature analysis for fault detection of induction machine–a review. 2021 4th Biennial International Conference on Nascent Technologies in Engineering (ICNTE),

Cardoso, A. J. M., Cruz, S. M. A., & Fonseca, D. S. B. (1997, 18-21 May 1997). Inter-turn stator winding fault diagnosis in three-phase induction motors, by park's vector approach. 1997 IEEE International Electric Machines and Drives Conference Record,

Cevasco, D., Koukoura, S., & Kolios, A. J. (2021). Reliability, availability, maintainability data review for the identification of trends in offshore wind energy applications. *Renewable and Sustainable Energy Reviews*, *136*, 110414. https://doi.org/10.1016/j.rser.2020.110414

Chen, B., Matthews, P. C., & Tavner, P. J. (2013). Wind turbine pitch faults prognosis using a-priori knowledge-based anfis. *Expert Systems with Applications*, *40*(17), 6863-6876.

Cho, S., Gao, Z., & Moan, T. (2016). Model-based fault detection of blade pitch system in floating wind turbines. Journal of Physics: Conference Series,

Choudhary, A., Mian, T., & Fatima, S. (2021). Convolutional neural network based bearing fault diagnosis of rotating machine using thermal images. *Measurement*, *176*, 109196. https://doi.org/10.1016/j.measurement.2021.109196

Erik Leandro, B., Levy Ely de Lacerda de, O., Jonas Guedes Borges da, S., Germano, L.-T., & Luiz Eduardo Borges da, S. (2012). Predictive maintenance by electrical signature analysis to induction motors. In Prof. Rui Esteves, A. (Ed.), *Induction motors* (pp. Ch. 20). IntechOpen. https://doi.org/10.5772/48045

Feng, Y., Tavner, P. J., & Long, H. (2010). Early experiences with uk round 1 offshore wind farms. *Proceedings of the Institution of Civil Engineers - Energy*, *163*(4), 167-181. https://doi.org/10.1680/ener.2010.163.4.167

Gecgel, O., Ekwaro-Osire, S., Gulbulak, U., & Morais, T. S. (2021). Deep convolutional neural network framework for diagnostics of planetary gearboxes under dynamic loading with feature-level data fusion. *Journal of Vibration and Acoustics*, *144*(3). https://doi.org/10.1115/1.4052364

Hammond, R., & Cooperman, A. (2022). *Windfarm operations and maintenance cost-benefit analysis tool (wombat)* (NREL/TP-5000-83712). https://www.nrel.gov/docs/fy23osti/83712.pdf

Jiang, Z., Han, Q., & Xu, X. (2020). Fault diagnosis of planetary gearbox based on motor current signal analysis. *Shock and Vibration*, *2020*, 8854776. https://doi.org/10.1155/2020/8854776

Junior, R. F. R., Areias, I. A. d. S., Campos, M. M., Teixeira, C. E., da Silva, L. E. B., & Gomes, G. F. (2022). Fault detection and diagnosis in electric motors using 1d convolutional neural networks with multi-channel vibration signals. *Measurement*, *190*, 110759. https://doi.org/10.1016/j.measurement.2022.110759

Kandukuri, S. T., Karimi, H. R., & Robbersmyr, K. G. (2016). Fault diagnostics for electrically operated pitch systems in offshore wind turbines. Journal of Physics: Conference Series,

Kandukuri, S. T., Senanayaka, J. S. L., Huynh, V. K., Karimi, H. R., & Robbersmyr, K. G. (2017). Current signature based fault diagnosis of field-oriented and direct torque–controlled induction motor drives. *Proceedings of the Institution of Mechanical Engineers, Part I: Journal of Systems and Control Engineering*, *231*(10), 849-866.

Kandukuri, S. T., Senanyaka, J. S. L., & Robbersmyr, K. G. (2019). A two-stage fault detection and classification scheme for electrical pitch drives in offshore wind farms using support vector machine. *IEEE Transactions on Industry Applications*, *55*(5), 5109-5118.

Khanjani, M., & Ezoji, M. (2021). Electrical fault detection in three-phase induction motor using deep network-based features of thermograms. *Measurement*, *173*, 108622. https://doi.org/10.1016/j.measurement.2020.108622

Kingma, D. P., & Ba, J. (2014). Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*.

Kumar, P., & Hati, A. S. (2022). Dilated convolutional neural network based model for bearing faults and broken rotor bar detection in squirrel cage induction motors. *Expert Systems with Applications*, *191*, 116290. https://doi.org/https://doi.org/10.1016/j.eswa.2021.116290

LeCun, Y., Boser, B., Denker, J., Henderson, D., Howard, R., Hubbard, W., & Jackel, L. (1989). Handwritten digit recognition with a back-propagation network. *Advances in neural information processing systems*, *2*.

Lee, J.-H., Pack, J.-H., & Lee, I.-S. (2019). Fault diagnosis of induction motor using convolutional neural network. *Applied Sciences*, *9*(15), 2950. https://www.mdpi.com/2076-3417/9/15/2950

Liu, Y., & Bazzi, A. M. (2017). A review and comparison of fault detection and diagnosis methods for squirrel-

cage induction motors: State of the art. *ISA transactions*, *70*, 400-409.

Lu, H., Pavan Nemani, V., Barzegar, V., Allen, C., Hu, C., Laflamme, S., Sarkar, S., & Zimmerman, A. T. (2023). A physics-informed feature weighting method for bearing fault diagnostics. *Mechanical Systems and Signal Processing*, *191*, 110171. https://doi.org/10.1016/j.ymssp.2023.110171

McKinnon, C., Carroll, J., McDonald, A., Koukoura, S., & Plumley, C. (2021). Investigation of isolation forest for wind turbine pitch system condition monitoring using scada data [Article]. *Energies*, *14*(20), 20, Article 6601. https://doi.org/10.3390/en14206601

Oppenheim, A. V. (1999). *Discrete-time signal processing*. Pearson Education India.

Park, J., Kim, C., Dinh, M. C., & Park, M. (2022). Design of a condition monitoring system for wind turbines [Article]. *Energies*, *15*(2), 16, Article 464. https://doi.org/10.3390/en15020464

Pfaffel, S., Faulstich, S., & Rohrig, K. (2017). Performance and reliability of wind turbines: A review. *Energies*, *10*(11), 1904. https://doi.org/10.3390/en10111904

Ren, Z., Verma, A. S., Li, Y., Teuwen, J. J. E., & Jiang, Z. (2021). Offshore wind turbine operations and maintenance: A state-of-the-art review. *Renewable and Sustainable Energy Reviews*, *144*, 110886. https://doi.org/10.1016/j.rser.2021.110886

Ruan, D., Wang, J., Yan, J., & Gühmann, C. (2023). Cnn parameter design based on fault signal analysis and its application in bearing fault diagnosis. *Advanced Engineering Informatics*, *55*, 101877. https://doi.org/https://doi.org/10.1016/j.aei.2023.101877

Sahraoui, M., Zouzou, S. E., Ghoggal, A., & Guedidi, S. (2010, 6-8 Sept. 2010). A new method to detect inter-turn short-circuit in induction motors. The XIX International Conference on Electrical Machines - ICEM 2010,

Santelo, T. N., de Oliveira, C. M. R., Maciel, C. D., & de A. Monteiro, J. R. B. (2022). Wind turbine failures review and trends. *Journal of Control, Automation and Electrical Systems*, *33*(2), 505-521. https://doi.org/10.1007/s40313-021-00789-8

Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*.

Singh, G. K., & Ahmed Saleh Al Kazzaz, S. a. (2003). Induction machine drive condition monitoring and diagnostic research—a survey. *Electric Power Systems Research*, *64*(2), 145-158. https://doi.org/https://doi.org/10.1016/S0378-7796(02)00172-4

Skowron, M., Orlowska-Kowalska, T., Wolkiewicz, M., & Kowalski, C. T. (2020). Convolutional neural network-based stator current data-driven incipient stator fault diagnosis of inverter-fed induction motor. *Energies*, *13*(6), 1475. https://www.mdpi.com/1996-1073/13/6/1475

Thorsen, O. V., & Dalva, M. (1995). A survey of faults on induction motors in offshore oil industry, petrochemical industry, gas terminals, and oil refineries. *IEEE Transactions on Industry Applications*, *31*(5), 1186-1196. https://doi.org/10.1109/28.464536

Trajin, B., Regnier, J., & Faucher, J. (2010). Comparison between vibration and stator current analysis for the detection of bearing faults in asynchronous drives. *IET electric power applications*, *4*(2), 90-100.

Walgern, J., Fischer, K., Hentschel, P., & Kolios, A. (2023). Reliability of electrical and hydraulic pitch systems in wind turbines based on field-data analysis. *Energy Reports*, *9*, 3273-3281. https://doi.org/10.1016/j.egyr.2023.02.007

Wang, X., Mao, D., & Li, X. (2021). Bearing fault diagnosis based on vibro-acoustic data fusion and 1d-cnn network. *Measurement*, *173*, 108518. https://doi.org/10.1016/j.measurement.2020.108518

Wei, L., Qian, Z., & Zareipour, H. (2019). Wind turbine pitch system condition monitoring and fault detection based on optimized relevance vector machine regression. *IEEE Transactions on Sustainable Energy*, *11*(4), 2326-2336.

Wilkinson, M., Hendriks, B., Spinato, F., Harman, K., Gomez, E., Bulacio, H., Roca, J., Tavner, P., Feng, Y., & Long, H. (2010). *Methodology and results of the reliawind reliability field study* European Wind Energy Conference , EWEC 2010, Warsaw, Poland. https://eprints.whiterose.ac.uk/83343/

Yuan, L., Lian, D., Kang, X., Chen, Y., & Zhai, K. (2020). Rolling bearing fault diagnosis based on convolutional neural network and support vector machine. *IEEE Access*, *8*, 137395-137406. https://doi.org/10.1109/ACCESS.2020.3012053

Zarei, J., & Poshtan, J. (2009). An advanced park's vectors approach for bearing fault detection. *Tribology International*, *42*(2), 213-219. https://doi.org/10.1016/j.triboint.2008.06.002

BIOGRAPHIES

Manuel S. Mathew is a PhD Research Fellow at the Information and Communication Technology department at the University of Agder, Norway. His interest is in the application of artificial intelligence in renewable energy systems particularly focusing on prognostics for wind farms. He completed his master's degree in Renewable Energy in 2021 from the University of Agder. In addition, he also holds a master's degree in systems engineering by research from the University of Brunei Darussalam. He did his bachelor's degree in electrical and electronics engineering from the Mahatma Gandhi University, India.

Surya Teja Kandukuri is a Senior Scientist at NORCE. He holds a part-time position as a researcher at University of Agder, Grimstad, Norway. He obtained his PhD in condition monitoring from the University of Agder in 2018. He has over 12 years of experience in industrial research within aerospace, energy, marine and oil & gas sectors, developing condition monitoring solutions for high-value assets. He received his master's degree in systems and control engineering from TU Delft, The Netherlands, in 2006 and bachelor's in electrical engineering from Nagarjuna University in India in 2003.

Christian W. Omlin has been a professor of Artificial Intelligence at the University of Agder since 2018. He has previously taught at the University of South Africa, University of the Witwatersrand, Middle East Technical University, University of the South Pacific, University of the Western Cape, and Stellenbosch University. His expertise is in deep learning with a focus on applications ranging from safety to security, industrial monitoring, renewable energy, banking, sign language translation, healthcare, bio conservation, and astronomy. He is particularly interested in the balance between the desire for autonomy using AI technologies and the necessity for accountability through AI imperatives such as explainability, privacy, security, ethics, and artificial morality for society's ultimate trust in and acceptance of AI. He received his Ph.D. from Rensselaer Polytechnic Institute and his MEng from the Swiss Federal Institute of Technology, Zurich, in 1995 and 1987, respectively.