Gaussian Process Kernel Transfer Enabled Method for Electric Machines Intelligent Faults Detection With Limited Samples

Chen, Jianjun; Hu, Weihao; Cao, Di; Zhang, Man; Huang, Qi; Chen, Zhe; Blaabjerg, Frede

Published in:
I E E E Transactions on Energy Conversion

DOI (link to publication from Publisher):
10.1109/TEC.2021.3070024

Publication date:
2021

Document Version
Accepted author manuscript, peer reviewed version

Citation for published version (APA):
Chen, J., Hu, W., Cao, D., Zhang, M., Huang, Q., Chen, Z., & Blaabjerg, F. (2021). Gaussian Process Kernel Transfer Enabled Method for Electric Machines Intelligent Faults Detection With Limited Samples. *I E E E Transactions on Energy Conversion*, *36*(4), 3481-3490. [9392260]. https://doi.org/10.1109/TEC.2021.3070024

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TEC.2021.3070024, IEEE Transactions on Energy Conversion

Gaussian Process Kernel Transfer Enabled Method for Electric Machines Intelligent Faults Detection with Limited Samples

Jianjun Chen, Weihao Hu, *Senior Member, IEEE,* Di Cao, Man Zhang, Qi Huang, *Senior Member, IEEE,* Zhe Chen, *Fellow, IEEE*, and Frede Blaabjerg, *Fellow, IEEE*

***Abstract*—Traditional Artificial Intelligence (AI) based fault detection approaches need a large amount of data for the model learning. However, in a real-world system, it is very difficult and expensive to obtain massive labeled fault data. In addition, the working conditions of a motor are usually variable, conventional fault diagnosis models with weak generalization ability can only be used for fault detection under constant working condition. The performance of traditional AI based approaches decreases when the working condition changes. To this end, a novel deep Gaussian process (GP) kernel transfer based few-shot learning method (RNGPT) is proposed in this paper for the fault detection of electric machines. First, a deep residual network (ResNet) is used to extract the features of the raw data. Then, the encoded latent feature vector is fed into the GP with kernel transfer ability to make the motor fault detection and classification. The proposed method uses much less data than the traditional AI based method to achieve fault diagnosis under variable working condition, and does not cause an overfitting problem. Experimental results of two case studies demonstrate that the proposed RNGPT model can accurately and effectively detect motor faults with limited labeled data under different working conditions. Experimental results of RNGPT with radial basis function (RBF) kernel model on simulation data present that the fault detection accuracy of the proposed method is about 16% higher than the conventional deep learning methods, 6% higher than other few-shot learning based methods in 5-shot and 4% higher in 1-shot. Finally, experimental on a real-world dataset, the RNGPT-RBF model still has the highest fault diagnosis accuracy in 5-shot (99.39 $\pm$ 0.09%) and 1-shot (98.55 $\pm$ 0.16%).**

***Index Terms*—Kernel transfer, Gaussian process, deep residual network, limited samples, electric machines, fault diagnosis.**

I. Introduction

ELECTRIC machines play an important role in industrial production, and their reliability is directly related to enterprise safety production and economic benefits [1], [2]. With the continuous development of science and technology, the application fields of electric machines have been widely used, such as electric vehicles, wind power generation, rail transit, etc. [3]. State monitoring and fault diagnosis of them can provide a reliable guarantee for the normal operation of industrial systems.

There are three main methods applied for electric machine faults diagnosis: The first method is based on the analytical model, which mainly uses the classical state estimation method to identify the fault, for instance, M. A. Mazzoletti et al. [4] proposed a model-based strategy for PMSM inter-turn short-circuit fault diagnosis. B. Aubert et al. [5] used an extended Kalman filter to estimate the number of faulty turns of a PMSM turn-to-turn short circuit, which is used to diagnose the turn-to-turn short circuit fault. In addition, torque modeling can also be used to analyze the electric machine state [6]. The second method is based on signal processing, which uses signal processing methods to analyze the selected signals to find out the characteristic quantities reflecting the motor faults. Common signal processing methods mainly focus on Fast Fourier Transform (FFT) [7], wavelet transform, Hilbert Transform (HT) [8], Empirical Mode Decomposition (EMD) [9], etc. Moreover, B. Wang et al. [10] and A. Sapena-Baño et al. [11] offered two harmonic order tracking analysis based methods that can also be used for electric machines fault diagnosis. Despite the previous two methods can be used for electric machine fault detection, the performance of these methods largely depends on the personal knowledge of signal processing and the ability of designers. The third method is based on artificial intelligence (AI) [12]-[14], which uses deep learning (DL) technology for feature extraction and classification of fault diagnosis [15], [16]. DL based faults diagnosis methods are a hot topic at present. The advantage of this method over the previous two methods is that there is no human intervention, which makes fault detection more efficient.

In recent years, numerous DL based methods have been proposed for electric machine faults detection due to its powerful ability of feature extraction and representation learning. The common DL based fault diagnosis methods include convolutional neural network (CNN), recurrent neural network (RNN) and deep Autoencoder etc. For instance, F. Wang et al. [17] recently proposed am improved cascade CNN method with progressive optimization for motor fault detection, which obtained a reasonable performance under non-stationary conditions. In addition, R. Liu et al. [18] proposed a multi-scale kernel based residual CNN model for electric machine fault detection, which avoid performance degradation when building

This work was supported by the Sichuan Science and Technology Program under Grants 2020JDJQ0037. (Corresponding author: Weihao Hu.)

Jianjun Chen, Weihao Hu, Di Cao, Man Zhang, Qi Huang are with the Power System Wide-area Measurement and Control Sichuan Province Key Laboratory, School of Mechanical and Electrical Engineering, University of Electronic Science and Technology of China Chengdu 611731, China (e-mail: chjj@std.uestc.edu.cn; whu@uestc.edu.cn; caodi@std.uestc.edu.cn; zhangman@uestc.edu.cn; hwong@uestc.edu.cn);

Zhe Chen and Frede Blaabjerg are with the Department of Energy Technology, Aalborg University, Aalborg, Denmark (e-mail: zch@et.aau.dk; fbl@et.aau.dk)

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TEC.2021.3070024, IEEE Transactions on Energy Conversion

a deeper network. Moreover, S. Shao et al. [19] proposed a deep CNN based approach for induction motor fault detection having multi-signals. Their model has a robust performance and accurate motor fault detection. For RNN based methods, J. Lei et al. [20] presented a wind turbine fault diagnosis model based on Long Short-term Memory (LSTM). Moreover, Z. An et al. [21] applied a RNN based model for motor bearing fault detection under time-varying working conditions and proposed a combination loss function, which can increase the training efficiency. For deep Autoencoder, it is also used in the field of fault diagnosis. For example, L. Wang et al. [22] applied deep Autoencoder for wind turbine state monitoring and achieved ideal results. In addition, F. Cheng et al. [23] proposed a model composed of a deep stacked Autoencoder and a support vector machine, which is used for DFIG wind turbine gearboxes fault diagnosis based on rotor current. In addition to the fault diagnosis methods described above, deep belief networks (DBNs) were also applied to electric machine fault detection [24].

Although the methods presented above have promising fault diagnosis performance, the conventional DL based electric machine fault diagnosis methods need a large amount of data for model training, therefore, these methods will be with less competitive performance in the case of limited samples. In the real world, it is hard and costly to obtain a large amount of labeled electric machines samples for model training. With limited training data, conventional deep architectures have the possibility to overfit the training data. Moreover, when the working conditions change, the performance of the model will also be affected. In order to address the described problems above, one of the solutions is to expand the original dataset. For instance, P. Liang et al. [25] applied Generative Adversarial Network (GAN) based method for fault diagnosis, GAN can generate fake samples with similar characteristics to real data. In addition, the original dataset can also be expanded by artificially adding noise, etc. Other solutions are using the transfer learning, for instance, S. Shao et al. [26] applied a transfer learning based approach for machine fault detection. However, this method needs a large amount of data for model pre-training.

Recently, few-shot learning approaches has gradually become a hot topic, which offers a promising approach to overcome the shortcomings of the fault diagnosis mentioned above. In general, few-shot learning aims to train a model under the condition of a limited labeled data. Face, speech and fingerprint recognition technologies, etc. can all be implemented by using few-shot learning methods. Moreover, few-shot learning uses different tasks similar to the target task to learn a learner [27]. A. Zhang et al. [28] proposed a siamese neural network based few-shot learning method for rolling bearing fault diagnosis and achieved promising results. In addition, S. Zhang et al. [29] and J. Wu et al. [30] also proposed two few-shot learning based methods for machine fault diagnosis. Inspired by these methods, in this paper, the “different tasks” can be regarded as the fault diagnosis tasks under different working conditions, and the “target task” can be regarded as a fault diagnosis task under an unseen working condition. As mentioned above, few-shot learning can not only address the problem of large amounts of data that are difficult to collect, but also solve the problem of applying the same model for fault diagnosis under different working conditions. Therefore, applying few-shot learning based method for electric machine faults detection seems to be an interesting approach.

In this paper, a novel deep GP kernel transfer based few-shot learning method (RNGPT) for electric machine fault detection is proposed. First, a deep Residual Network (ResNet) [31] is applied to feature extraction, which improves the training efficiency and increases the fault detection accuracy. Then a Gaussian process (GP) [32] with kernel transfer capability is applied to process the features extracted by ResNet and classify the health state of electric machine. In this paper, it is different from the conventional transfer learning that GP kernel transfer uses the identical kernel for all fault diagnosis tasks even though under different working conditions. Therefore, when the fault diagnosis task is carried out under an unseen working condition, only the posterior distribution needs to be recalculated and there is no need to update the model parameters or make fine-tuning. Finally, the proposed model will be tested on the dataset of an unseen working condition. The contributions of this paper can be summarized as follows:

1) The proposed method can achieve high classification accuracy utilizing only few-shot training samples. It distinguishes from traditional DL methods that require a large amount of labeled data, which are difficult and costly to obtain in practice.
2) More comprehensive electric machine working conditions are considered in this paper. In addition, the training, validation and test of the model are carried out under different working conditions. Experimental results demonstrate that the proposed RNGPT model has a stronger generalization ability.
3) In most cases, the proposed model has the best performance compared with other state-of-the-art methods. The RNGPT can address the problems where the traditional DL based methods need large amounts of data for training. Compared with conventional intelligent fault detection models, the proposed model does not need to be retrained or fine-tuned when the working conditions of the electric machine are changed.

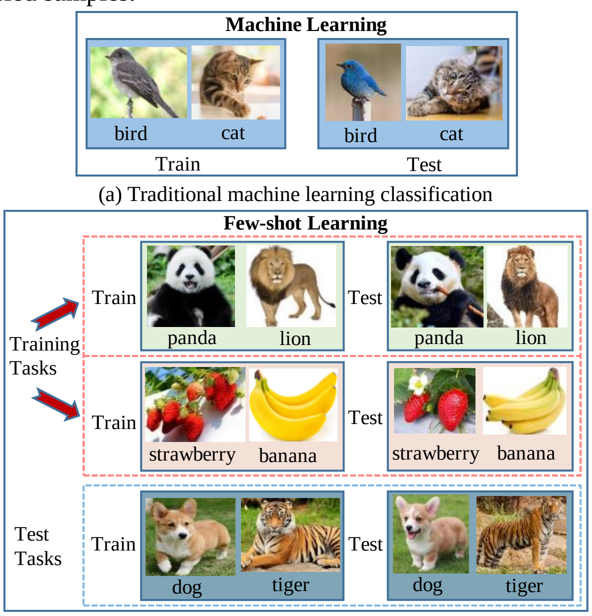
The remainder of this paper is organized as follows. Section II includes the fundamental methodology of the RNGPT model. The framework of the proposed electric machine fault detection model is introduced in Section III. The case study and results analysis are presented in Section IV. Finally, the conclusions are given in Section V.

II. Methodology and Proposed Framework

This section mainly introduces the basic methodology and the framework of the proposed method. The methodology of a few-shot learning is presented in Section II-A. The descriptions of the proposed framework are represented in Section II-B.

A. *Few-shot Learning*

Traditional machine learning technology needs a large-scale data to train a model. For example, a study on the classification of MNIST, there are 10 classes and 6000 samples in the training set. On average, there are about 600 samples per class. In some practical tasks, machine learning methods face the case of "small data" set. For example, image data for some rare species are difficult to be collected with very few labeled images. In this case, the neural network model needs to be trained based on a small number of sampling data. Therefore, it is important to develop a reliable model for a task with only a few labeled training samples. Few-shot learning aims to train a model to identify unseen targets during training with limited labeled examples [33], which could address the previous problem. Few-shot learning is the application of meta learning in supervised learning, in the case of small sample classification, which is defined as a task with N-class and K-sample (N-way K-shot). In the training phase of few-shot learning, it has many different classification tasks where the purpose is to train a model that can be transferred to a new class with only a few labeled samples.



(a) Traditional machine learning classification

(b) Few-shot learning classification

Fig. 1. The difference between traditional machine learning method and few-shot learning method. (a) Traditional machine learning method has only one task for model training, (b) Few-shot learning has multiple classification tasks for model training.

Let $D_s = \{(x_k, y_k)\}_{k=1}^{K}$ (seen class) be a support set (referred to in conventional deep learning as a training set), which contains input-output pairs. When k is equal to one, this method is called one-shot learning. Let $D_q = \{(x_t, y_t)\}_{t=1}^{T}$ (unseen class) be a query set (referred to in conventional deep learning as a test set), typically T is one order of magnitude greater than K. Let $T = \{D_s, D_q\}$ denote the group of the support and query sets in a training task, $C = \{T_m\}_{m=1}^{M}$ denote a collection of all these tasks. Few-shot learning models are trained on a series of random tasks which are sampled from collection C. In the test task, a new task $T_* = \{D_s^*, D_q^*\}$ is sampled from a test set (which differs from conventional deep learning test set) for model performance validation. The purpose of the test task is to fine-tune the model based on the support set D_s^*, and then make category predictions on the samples of the query set D_q^*. For a more intuitive comprehension of the difference between few-shot learning and traditional machine learning, a visualized example is presented in Fig. 1.

B. *Architecture of the Proposed Network*

The framework of the proposed method is introduced in this section. To establish a strong generalization model for electric machine intelligent faults detection with few signal samples, inspired by the few-shot learning, a GP based kernel transfer method is applied in this paper. The overall framework of the proposed method is illustrated in Fig. 2. Each task can be regarded as a fault diagnosis task under a specific working condition. It can be observed in Fig. 2 that a sub-sampling window containing a series of raw data points is collected to be one sample, and there is a sampling interval between each sub-sampling window. A dataset for model training, validation and testing is a combination of all these samples with different EM health states. The representation layer extracts the discriminative features of the input data (three-phase current and vibration signal). Then the encoded feature maps are fed into a decision making layer to perform the electric machine health state classification. Finally, the trained base model, which is called feature fusion model will be transferred to different electric machine working conditions data for faults diagnosis. The difference between conventional transfer learning methods is that the proposed method in this paper uses the identical kernel for all faults diagnosis tasks which means that the fault diagnosis tasks under different working conditions can share a set of "multi-function" hyper parameters. Therefore, there is no additional encoding module required.

1) Representation Layer: Since the characteristics of electric machine signals are different under different fault types, it is reasonable to detect the faults by analyzing the input inner features. And the discriminative features extracting from raw signal is playing a crucial role to improve the GP joint distribution approximation, which has a great influence on the faults detection ability of the model. Deep neural network is known for its powerful feature extraction ability [34], which is a machine learning model composed of several stacked nonlinear processing layers and has made bright achievements in the field of smart grid applications. However, as the depth of neural network increases, the problem of gradient explosion or gradient vanishing will be caused, which makes the solver difficult to obtain the optimal solution and makes the neural network model to have a higher training error. The advantages of deep ResNet is that it could not only increase the learning efficiency of deeper network, but also increase the accuracy of the model. Moreover, the residual network can avoid the gradient vanishing or gradient explosion problem described previously with the increase of network layers, which makes it possible to develop a very deep neural network. In this paper, a deep ResNet is applied to extract the raw data features. These encoded features will boost the faults diagnosis performance of decision making layer.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TEC.2021.3070024, IEEE Transactions on Energy Conversion

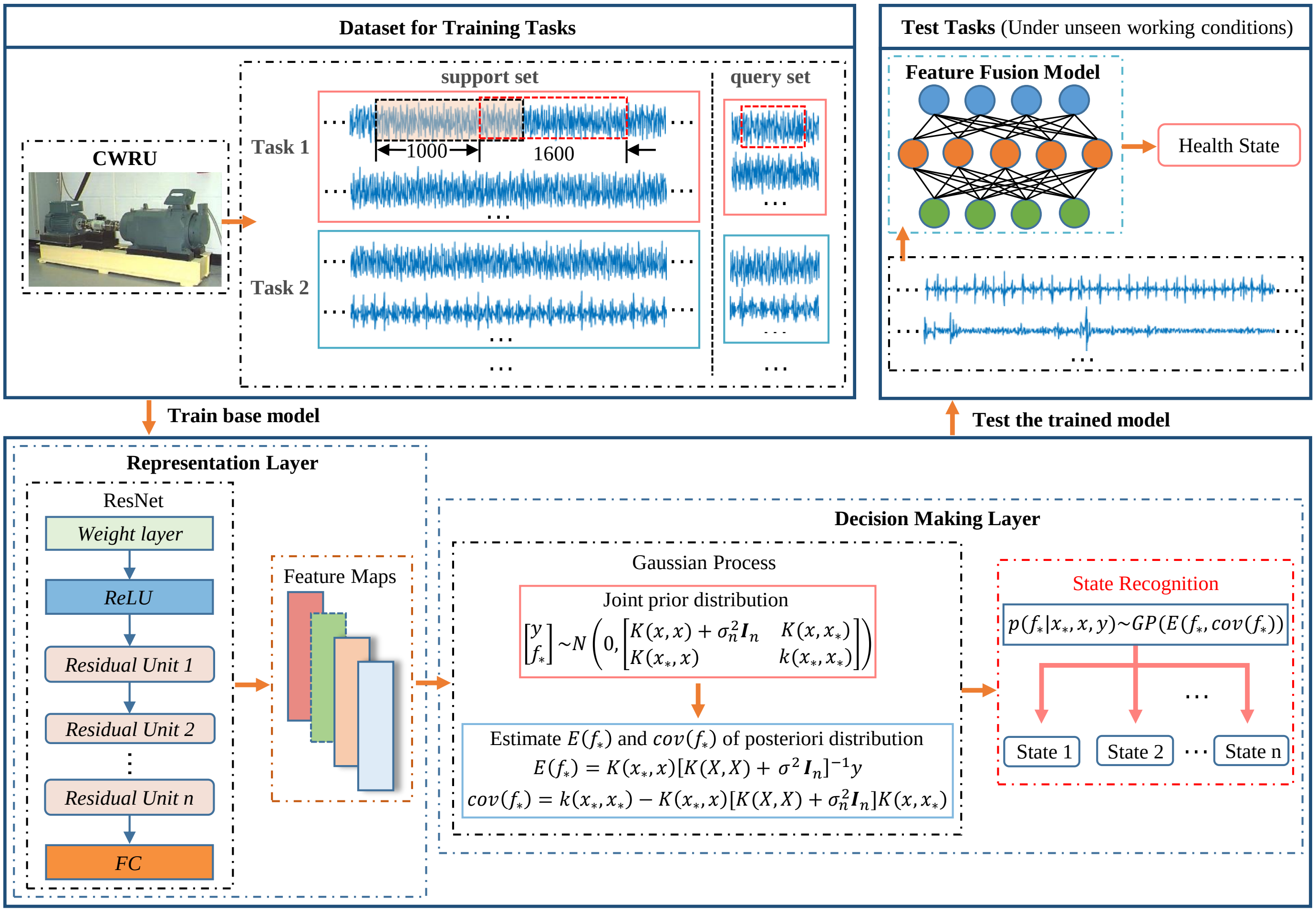


Fig. 2. The overall framework of the proposed method for electric machine intelligent faults detection. The few-shot learning based model uses the current and vibration signal as inputs, extracts the discriminative features, and calculates the GP joint distribution to recognizing the electric machine state.

In the representation layer of Fig. 2, a conventional neural network based ResNet with 10 stacked residual units (ResNet-10) is applied to process the electric machine signals. The difference between a common network and a residual unit is shown in Fig. 3.

It shows in Fig. 3 that the output $M(x)$ of a common neural network is a desired mapping of input x, which can be described as follows:

$$M(x) = f(G(x, W)), \tag{1}$$

where $G(x, W)$ is a mapping function about input x and weight matrix W, $f(\cdot)$ is the activation function $ReLU$. Assuming that x_i is the input of the i-th residual unit, the output can be described as follows:

$$x_{i+1} = f(F(x_i,\ W_i) + x_i), \tag{2}$$

where the $F(x_i,\ W_i)$ is the residual function mapping, W_i is the corresponding weight matrix. It should be noted that the dimensions of x_i and $F(x_i,\ W_i)$ should be equal, otherwise, it could then perform a linear mapping W_s to match the two dimensions

$$x_{i+1} = f(F(x_i,\ W_i) +\ W_s x_i), \tag{3}$$

The form of residual function $F(\cdot)$ is flexible and changeable. In addition to stack two convolution layers, it can also stack three convolution layers.

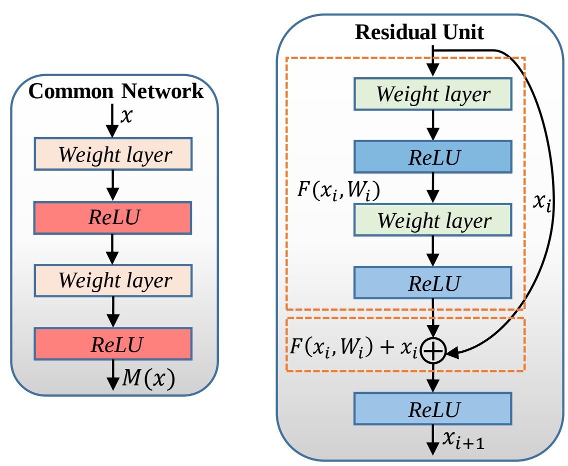
Compared with conventional CNN fitting the underlying mapping directly, the ResNet fits a residual between two adjacent signals, which simplifies the training process of the model. Assuming the underlying feature non-linear mapping is

$$H(x) = F(x_i,\ W_i) + x_i, \tag{4}$$

the ResNet needs to fit a function mapping of

$$F(x_i,\ W_i) = H(x) - x_i, \tag{5}$$

The results show that the model applying ResNet is easier to train and has better performance. The experimental results will be presented in Section III. After processed by the



(a) (b)

Fig. 3. The difference between a common neural network and a residual unit. (a) A common neural network, (b) A residual unit of ResNet.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TEC.2021.3070024, IEEE Transactions on Energy Conversion

representation layer, the raw data features will be passed to the decision making layer for further processing.

2) Decision Making Layer: The decision making layer is a GP based electric machine health state classifier. It shows in Fig. 2 that the input of this layer is the non-linear mapping based on Eq. (4): $H(x) \rightarrow \boldsymbol{h}$, where $\boldsymbol{h}$ is a desired feature map.

A GP can be determined by a mean function $\mu(x)$ and a covariance function (kernel function) $k(x,\, x')$ uniquely. According to the definition of GP, the distribution functions f can be denoted by

$$f(x) \sim GP(\mu(x), k(x,\, x')), \tag{6}$$

where the $\mu(x)$ and $k(x,\, x')$ can be calculated by

$$\mu(x) = E(f(x)), \tag{7}$$

$$k(x,\, x') = E[\{f(x) - m(x)\}\{f(x') - m(x')\}], \tag{8}$$

In general, the value of the mean function $\mu(x)$ is assumed to 0. This is because in the application of GP in machine learning, the data preprocessing method, zero-mean normalization is often necessary if we do not have any prior knowledge. The covariance function $k(x,\, x')$ can capture the relationship between different input points in the input space.

In practical application, the Gaussian noise should also be considered:

$$\varepsilon \sim N(0, \sigma_n^2), \tag{9}$$

where ε is the Gaussian noise with variance σ_n^2, and it is completely independent for the $f(x)$. Based on Bayesian probability theory, by establishing a priori distribution function on a given training dataset $D = \{(x_n, y_n)\}_{n=1}^{N}$, the output distribution y_n with Gaussian noise can be obtained as follows:

$$y_n = f(x_n) + \varepsilon \sim N(0, K(X, X) + \sigma_n^2 \boldsymbol{I}_n), \tag{10}$$

where $\boldsymbol{I}_n$ is an $N \times N$ identity matrix; $K(X, X)$ is a Gram matrix:

$$K(X, X) = (K_{ij})_{n \times n} = k(x_i, x_j), i, j = 1, 2, \cdots n, \tag{11}$$

The joint prior distribution of observations and predictions can be obtained as follows:

$$\begin{bmatrix} y \\ f_* \end{bmatrix} \sim N\left(0, \begin{bmatrix} K(x, x) + \sigma_n^2 \boldsymbol{I}_n & K(x, x_*) \\ K(x_*, x) & k(x_*, x_*) \end{bmatrix}\right), \tag{12}$$

where y_* is the predicted value; x_* is a testing point, which has the same Gaussian distribution with the input points x in training set D. $K(x, x_*) = K(x_*, x)^T$ is the $n \times 1$ covariance matrix between test point x_* and input points x. $k(x_*, x_*)$ is the covariance of point x_*. According to the previous joint prior distribution, the posteriori distribution probability of f_* can be obtained by Bayes' rule as follows:

$$p(f_* | x_*, x, y) \sim GP(E(f_*), cov(f_*)), \tag{13}$$

where the mean $E(f_*)$ and covariance $cov(f_*)$ of the prediction point f_* can be obtained as follows:

$$E(f_*) = K(x_*, x)[K(X, X) + \sigma_n^2 \boldsymbol{I}_n]^{-1} y, \tag{14}$$

$$cov(f_*) = k(x_*, x_*) - K(x_*, x)[K(X, X) + \sigma_n^2 \boldsymbol{I}_n] K(x, x_*), \tag{15}$$

The mean $E(f_*)$ can be seen as the estimation value of the prediction point f_* and the $cov(f_*)$ can reflect the reliability of this value.

Decision making layer calculates first the joint prior distribution of the underlying features of training and test set, the specific calculation process is based on Eq. (12). After obtaining the prior distribution, the proposed model calculates the probability corresponding to each electric machine health state according to Eq. (13). The specific parameters in Eq. (13) can be calculated by combining Eqs. (14) and (15). The state which corresponds to the highest probability is the motor state predicted by the model.

C. Training of The Proposed Model

In this paper, the GP hyper-parameters of the proposed model are determined by the maximum likelihood estimation method. In addition, the GP sample function of the model in this paper is disconnected from the hyper-parameters, which need to be determined. So that, the proposed model simply focuses on exploring the hyper-parameters of a GP kernel that has the maximum marginal likelihood (marginalizing out the GP sample function). With this approach, it means that it is reasonable to calculate the posterior distribution probability $P(y_t^* | x, \theta, \emptyset)$ directly and skip a complicated intermediate reasoning step. Let y_t^* denotes the target data, D_x denotes the input data (a collection of support set and query set). The marginal likelihood function $L(\theta, \emptyset)$ can be described as

$$L(\theta, \emptyset) = \log P(y_t^* | D_x, \theta, \emptyset) = \frac{1}{2}(y_t^*)^T C^{-1} y_t^* + \frac{1}{2}\log|C| + \alpha, \tag{16}$$

where α is a constant. C is the kernel between all task inputs and this is related to the estimation of hyper-parameter θ and neural network weight $\emptyset$. The proposed model is optimized by maximizing the marginal likelihood function.

Algorithm 1 shows the training and test process of the proposed faults detection model.

Algorithm 1 Training and test procedures of the proposed model

	Training tasks
Input:	Current signal and vibration signal
Output:	The health state of electric machine
Setting:	Parameters of few-shot learning: n and m (n-way, m-shot)
Initialize:	Randomly initialize the ResNet-10's parameters $\emptyset$
1:	**for** epoch=1:stop_epoch **do**
2:	Randomly sample training task $T = \{D_s, D_q\} \sim C$
3:	Get x and y from D_s, get x_* from D_q, then predict signal f_* in D_q
4:	Feature extracting: $H(x) \rightarrow \boldsymbol{h}$ (Eq. (4))
5	Estimate the posteriori distribution probability: $p(f_* \| T, \theta, \emptyset)$ (Eq. (13) ← $E(f_*)$ and $cov(f_*)$ (Eq.(14) and Eq.(15))
6:	Calculate the loss function: $f_{loss} = -L(\theta, \emptyset)$ (Eq. (16))
7:	Update parameters of GP: $\theta_{t+1} = \theta_t - \beta \nabla_{\theta_t} f_{loss}$
8:	Update parameters of ResNet-10: $\emptyset_{t+1} = \emptyset_t - \gamma \nabla_{\emptyset_t} f_{loss}$
9	**end for**
	Test tasks
10	Randomly sample test task $T_* = \{D_s^*, D_q^*\}$
11	Get x and y from D_s^*, get x_* from D_q^*, then predict clean signal f_* in D_q^*
12	Estimate posteriori distribution probability of f_* : $p(f_* \| x_*, x, y)$ (Eq. (13)) ← $E(f_*)$ and $cov(f_*)$ (Eq.(14) and Eq.(15))
13	**end**

The parameters of the ResNet-10 are denoted as $\emptyset = \{W, b\}$, where $W = \{W_i, W_1, W_2, \cdots, W_{10}\}$ represent the weight parameters of the ResNet-10, and $b = \{b_i, b_1, b_2, \cdots, b_{10}\}$ represent the bias parameters of the ResNet-10. The parameters of the GP are denoted as $\theta = \{\boldsymbol{K}, \sigma_n^2 \boldsymbol{I}_n\}$, where $\boldsymbol{K}$ represents

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TEC.2021.3070024, IEEE Transactions on Energy Conversion

the covariance matrix and $\sigma_n^2 \boldsymbol{I}_n$ represents the noise.

Two stages are included in Algorithm 1: In the training phase, the input of the model is motor current or vibration signal. The data are processed by a neural network and then sent to GP to perform motor states recognition. The loss function of the proposed model can be derived based on Eq. (16) as

$$f_{loss} = -L(\theta, \emptyset) = -\log P(y_t^* | D_x, \theta, \emptyset), \quad (17)$$

In this paper, the gradient descent method is applied to update the parameters to minimize the loss function:

$$\theta_{t+1} = \theta_t - \beta \nabla_{\theta_t} f_{loss}, \quad (18)$$

$$\emptyset_{t+1} = \emptyset_t - \gamma \nabla_{\emptyset_t} f_{loss}, \quad (19)$$

where β and γ are the learning rate, $\beta \nabla_{\theta_t} f_{loss}$ and $\gamma \nabla_{\emptyset_t} f_{loss}$ represent the gradient of the loss function.

In the test phase, the input of the model is also the motor current or vibration signal, but its working conditions may be different from those of the conditions in the training phase. In this phase, even if the input never appears in the training phase, the parameters of the fault diagnosis model are not updated anymore. This is because the model uses the identical kernel for all faults diagnosis tasks, even if the input data comes from different working conditions, the model parameters can be transferred.

III. Results Analysis

In this section, the performances of the proposed model with different kernels (RNGPT-BNCosSim, RNGPT-CosSim and RNGPT-RBF) will be evaluated in two case studies. Accuracy is used as the evaluation criteria of fault diagnosis performance. The case study I is presented in Section IV A, where the proposed model is applied to simulation dataset in order to demonstrate its feasibility in fault diagnosis. Then the model is tested on a real-world dataset from Case Western Reserve University Bearing Data Center Website [35], the detailed analysis is presented in Section IV B. Several state-of-the-art few-shot learning algorithms for comparison are also presented in this paper: MAML [36], ProtoNet [37] and Baseline [38]. Moreover, two conventional DL methods WDCNN [39] and Capsule-Net [40] are also used to compare with the proposed method.

TABLE I
Detailed Hyper-Parameters of the Proposed Methods

Hyper-parameters	RNGPT-BNCosSim	RNGPT-CosSim	RNGPT-RBF
Learning rate	10^{-3}	10^{-3}	10^{-3}
Number of layers	10	10	10
Kind of layers	ResNet (CNN)	ResNet (CNN)	ResNet (CNN)
Epoch	400	400	400
Mini-batch	16	16	16
Optimizer	Adam	Adam	Adam
Loss function	Marginal likelihood	Marginal likelihood	Marginal likelihood
Activation function	ReLU	ReLU	ReLU
GP Kernel function	BNCosSim	CosSim	RBF

In this paper, the learning rate for GP based decision making layer is $\beta = 10^{-4}$ and for ResNet based feature extractor is $\gamma = 10^{-3}$. All parameters are optimized and updated by an Adam optimizer. The other detailed parameters of the experiment are introduced in the following subsections. The training process of the proposed model are completed on a 64-bit computer with one Intel(R) Core(TM) i7-9800X CPU of 3.8 GHz and one RTX2080 Ti GPU of 11 GB VRAM, besides, it also has one 32.0 GB RAM. The code of proposed model is written in Pycharm platform with Pytarch of GPU version. The values of the detailed hyper-parameters of the proposed methods are summarized in Table I. All these hyper-parameters are tuned and determined by Grid Search method [41].

A. Case Study I: Simulation Dataset

In this case study, the dataset used for the motor faults diagnosis experiment is obtained by simulation. The proposed model in this research is trained to identify different motor health states with only a small number of training samples. A three-phase asynchronous induction motor is simulated under three different sizes of loads (working conditions) by ANSYS Electronics Suite software [42]. The circuit electromagnetic performance and system design can be simulated by this software uniquely. Moreover, it can also evaluate other important mechanical effects, such as temperature, vibration etc. [42]. The detailed parameters of the motor are summarized in Table II. The three-phase current signal is collected and used as the input of the proposed fault diagnosis model. Each working condition contains seven motor health states, therefore, a total of 21 (3×7) sets of data are obtained in the simulations. The descriptions of each motor health state with three different sizes of loads are shown in Table III.

TABLE II
Detailed Parameters of Motor

Parameter	Setting Value
Rated power	11 kW
Rated speed	1458 rpm
Rated load torque	75 Nm
Rated voltage	380 V
Rated frequency	50 Hz
Efficiency & power factor	0.85 and 0.86

TABLE III
Descriptions of Motor Health State

Motor Load (Nm)	Health State	Description
	Normal	Healthy motor without fault
	BRB	There are 4 Broken squirrel-cage Bars, two adjacent, and the other two adjacent in the opposite pole direction
	ISC	Inter-turn Short Circuit of a slot in phase A
68	LM	Sudden Loss of load
70	SPF	Stator Phase Failure
72	RDE	Rotor Dynamic Eccentricity, the rotor rotates once, the trajectory of the rotor center forms a circle with a radius of eccentric displacement
	RSE	Rotor Static Eccentricity, the shaft of the rotor is fixed, and the center of the trajectory of the rotor rotates at one point

It can be observed from Table III that the motor with 68, 70, and 72 Nm loads are simulated and the data with seven types of motor health states are obtained. In order to construct a dataset with only a few samples which is used to verify the performance of the model. The running time of motor simulation is four seconds, and the data sampling frequency is 1 kHz. Therefore, the length of the data under each health state is 4000. It is shown in Fig. 2 that a sub-sampling window containing 200 data points is collected to be one sample, and with 50 data points between each window. As described above, the sample size of each health state is 77. In total, there are 539 set of samples under each working condition for model training, validation and testing.

1) Process of Model Training: In order to explore the fault diagnosis model with high accuracy and strong generalization under variable working conditions. A dataset with 70 Nm load is selected for model training, 68 Nm load for validation and 72 Nm load for model final performance testing. The proposed faults detection model is trained for 400 epochs to get the highest accuracy. During training phase, the parameters of the model with the highest accuracy in the validation set will also be collected at the same time in each epoch. Therefore, the performance of the model can be guaranteed under different working conditions. All the tasks are 5-ways (5 of the 7 types of health state are randomly selected) with 1-shot or 5-shot (1 or 5 health state samples are randomly selected in the support set). Otherwise, there are 12 samples included in each query set. In the representation layer, a 256-dimension feature vector is extracted from each three-phase current signal sample. This feature map is passed to the input layer of a decision making layer in order to make a 5-class classification (5-way).

2) Evaluation of Model Performance: The health states detection performances of the various methods are presented in Table IV.

TABLE IV
PERFORMANCE OF DIFFERENT MODEL IN SIMULATION DATASET

Health States Detection Model	**Health States Detection Accuracy (100%)**	
	5-shot	**1-shot**
MAML	89.88 ± 0.41	88.84 ± 1.43
ProtoNet	90.51 ± 1.24	90.28 ± 0.70
Baseline	89.60 ± 0.67	89.19 ± 1.45
RNGPT-BNCosSim	90.68 ± 0.70	90.42 ± 0.68
RNGPT-CosSim	90.35 ± 0.77	90.17 ± 0.74
RNGPT-RBF	96.10 ± 0.29	93.28 ± 0.68
Conventional method	**Health States Detection Accuracy**	
WDCNN	80.23 ± 0.82	
Capsule-Net	84.20 ± 1.76	

It shows that the proposed RNGPT model has the best fault detection performance compared with other methods on the simulation dataset. RNGPT with RBF kernel has the highest detection accuracy in 5-shot ($96.10 \pm 0.29\%$) and 1-shot ($93.28 \pm 0.68\%$). Compared with RNGPT with BNCosSim and CosSim kernel, RNGPT-RBF outperforms these two models and has at least 5% more accurate in 5-shot and 2% more accurate in 1-shot. And compared with three few-shot learning based methods, the proposed RNGPT model can detect the health states of the motor more accurately, except RNGPT-CosSim model. This is because the performance of the model is degraded by the CosSim kernel function. Therefore, an appropriate GP kernel can make the proposed model more effective. In addition, compared with Capsule-Net, the detection accuracy of RNGPT-RBF is about 12% higher. There are two main reasons for this result: one is the limited data for model learning, which makes Capsule-Net overfitting; another is that is that when the working condition changes during the model test, the traditional DL method has a weak generalization ability and as a consequence, the detection accuracy of this model will decrease if it does not fine-tune the model parameters.

B. Case Study II: Case Western Reserve University Bearing Dataset

In this case study, the dataset used for the motor faults diagnosis experiment is obtained from Case Western Reserve University (CWRU) Bearing Data Center. The detailed descriptions of the fault type and diameter are shown in Table V. It shows that there are three different type of motor bearing faults, including inner raceway, rolling element and outer raceway. Each type of fault has three different sizes, which are 0.007, 0.014 and 0.021 inches. Therefore, there are totally 10 kinds of motor bearing health states for each working condition, including 9 kinds of fault states and one kind of normal state. The vibration signal is collected and used as the input of the proposed fault diagnosis model, the data sampling frequency is 12 kHz. In order to construct a small-scale dataset, a sub-sampling window containing 1600 data points is collected to be one sample every 1000 points, as shown in Fig. 2. In total, there are 1000 set of samples under each working condition for model training, validation and testing.

TABLE V
DETAILED PARAMETERS OF MOTOR FOR CASE STUDY II

Motor Load (hp)	**Fault Diameter (inches)**	**Fault Type**		
1	0.007	Inner Raceway	Rolling Element	Outer Raceway
	0.014			
	0.021			
2	0.007			
	0.014			
	0.021			
3	0.007			
	0.014			
	0.021			

1) Process of Model Training:

In this research, vibration data of drive end bearing fault with three different type of loads (1-3 hp) are selected for model training, validation and testing. The dataset with 1 hp load is selected for model training, 2 hp load for validation and 3 hp load for model final performance testing. The proposed faults detection model is trained for 400 epochs to get the highest accuracy. During the training phase, the parameters of the model with the highest accuracy in the validation set will also be collected at the same time in each epoch. Therefore, the performance of the model can be guaranteed under different working conditions. All the tasks are 5-way (5 of the 10 types of health state are randomly selected) with 1-shot or 5-shot (1 or 5 health state samples are randomly selected in the support set). And there are 16 samples included in each query set. In the representation layer, a 512-dimension feature vector is extracted from the vibration signal. This feature map is passed to the input layer of decision making layer then to make a 5-class classification (5-way).

2) Evaluation of Model Performance:

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TEC.2021.3070024, IEEE Transactions on Energy Conversion

The evolution of the loss function value over 400 epochs is illustrated in Fig. 4. It shows that with the continuous optimization of the model, the value of the loss function is constantly decreasing, which demonstrates that the model proposed in this article can be applied to real-world data.

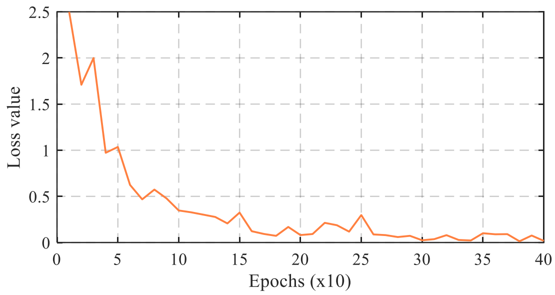
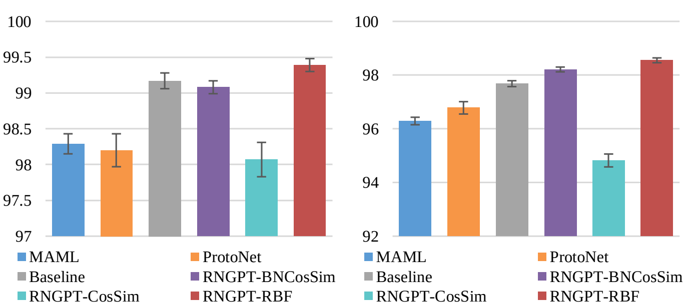


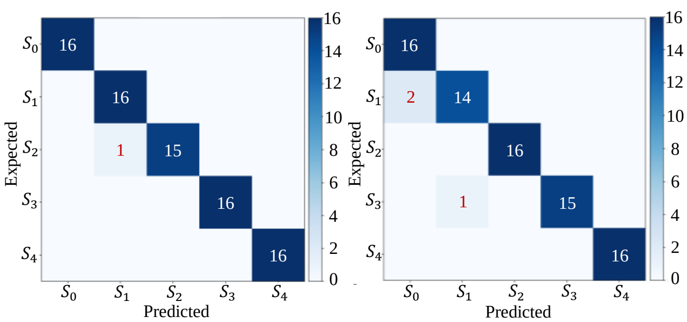
Fig. 4. The evolution of the loss function value over 400 Epochs.

The health states detection performances of different methods are presented in Fig. 5. It can be observed in Fig. 5 that the proposed RNGPT-RBF model has the best fault detection performance compared with other methods on the CWRU bearing dataset. RNGPT with RBF kernel has the highest fault diagnosis accuracy in 5-shot (99.39 $\pm$ 0.09%) and 1-shot (98.55 $\pm$ 0.16%). In addition, the method proposed in this paper has smaller variance, which indicates that it has a more stable fault diagnosis performance.



(a) 5-shot (b) 1-shot

Fig. 5. The faults detection accuracy of different models in CWRU dataset: (a) 5-shot, (b) 1-shot.



(a) 5-shot (b) 1-shot

Fig. 6. The confusion matrices of RNGPT-RBF model health state detection results: (a) 5-shot, (b) 1-shot.

The confusion matrices of RNGPT-RBF in the testing phase are presented in Fig. 6. S_0 to S_4 represent 5 randomly sampled motor health states. The abscissa represents the motor health states detected by the model, and the ordinate represents the actual states. Fig. 6 (a) is the confusion matrix of RNGPT-RBF in 5-shot scenario and it demonstrates that the proposed model can correctly detect the motor states with low error. Fig. 6 (b) is the confusion matrix of RNGPT-RBF in 1-shot scenario and it can be seen that the number of misjudgments is more than that shown in Fig. 6 (a). The reason is that, compared with 5 input samples in the 5-shot scenario, only one sample is input into the model in 1-shot scenario.

In summary, the results in Section-A and B demonstrate that the proposed RNGPT model can accurately and effectively detect motor health states with limited labeled data under different working conditions.

IV. Conclusion

A novel GP kernel transfer based few-shot learning method is proposed in this paper for electric machine fault diagnosis under variable working conditions. The diagnostic knowledge learns from limited current signal and vibration signal. The trained model can be transferred to other unseen working conditions without parameters updating and fine-tuning. Two case studies are carried using two different electric motor health state datasets, including a simulation dataset obtained by *ANSYS* software and a real-world bearing dataset from CWRU. In the first case, the proposed model has achieved the highest fault detection accuracy of 96.10 $\pm$ 0.29% in 5-shot and 93.28 $\pm$ 0.68% in 1-shot. In the second case, the proposed model also has the best fault detection performance in 5-shot and 1-shot, the detection accuracy is 99.39 $\pm$ 0.09% and 98.55 $\pm$ 0.31% respectively. Comparative case studies have shown that the proposed approach can achieve the state-of-the-art performance. The results demonstrate that, when the limited samples of the electric machine are available and the working conditions are variable, it is possible and significant to explore a fault detection model to exploit the underlying shared features for diagnostics.

The detection of early and hidden failures is still a challenge issue which needs to be addressed. Future research will focus on the early fault diagnosis and weak fault diagnosis of electric machine.

References

[1] X. Sun, Z. Shi, G. Lei, Y. Guo, and J. Zhu, “Multi-objective design optimization of an IPMSM based on multilevel strategy,” *IEEE Trans. Ind. Electron.*, vol. 68, no. 1, pp. 139-148, Jan. 2021.

[2] J. Hang, S. Ding, J. Zhang, M. Cheng, W. Chen and Q. Wang, "Detection of interturn short-circuit fault for PMSM with simple fault indicator," *IEEE Trans. Energy Convers.*, vol. 31, no. 4, pp. 1697-1699, Dec. 2016.

[3] X. Sun, Z. Jin, Y. Cai, Z. Yang, and L. Chen, “Grey wolf optimization algorithm based state feedback control for a bearingless permanent magnet synchronous machine,” IEEE Trans. Power Electron., vol. 35, no. 12, pp. 13631-13640, Dec. 2020.

[4] M. A. Mazzoletti, G. R. Bossio, C. H. De Angelo and D. R. Espinoza-Trejo, "A Model-Based Strategy for Interturn Short-Circuit Fault Diagnosis in PMSM," *IEEE Trans. Ind. Electron.*, vol. 64, no. 9, pp. 7218-7228, Sept. 2017.

[5] B. Aubert, J. Régnier, S. Caux and D. Alejo, "Kalman-Filter-Based Indicator for Online Interturn Short Circuits Detection in Permanent-Magnet Synchronous Generators," *IEEE Trans. Ind. Electron.*, vol. 62, no. 3, pp. 1921-1930, March 2015.

[6] X. Sun, J. Wu, G. Lei, Y. Cai, X. Chen, and Y. Guo, “Torque modeling of a segmented-rotor SRM using maximum-correntropy-criterion-based LSSVR for torque calculation of EVs,” *IEEE J. Emerg. Sel. Topics Power Electron.*, to be published, DOI: 10.1109/JESTPE.2020.2977957.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TEC.2021.3070024, IEEE Transactions on Energy Conversion

[7]E. Cabal-Yepez, A. G. Garcia-Ramirez, R. J. Romero-Troncoso, A. Garcia-Perez and R. A. Osornio-Rios, "Reconfigurable Monitoring System for Time-Frequency Analysis on Industrial Equipment Through STFT and DWT," *IEEE Trans. Ind. Inform.*, vol. 9, no. 2, pp. 760-771, May 2013.

[8]B. Xu, L. Sun, L. Xu and G. Xu, "Improvement of the Hilbert Method via ESPRIT for Detecting Rotor Fault in Induction Motors at Low Slip," *IEEE Trans. Energy Convers.*, vol. 28, no. 1, pp. 225-233, March 2013.

[9]R. Valles-Novo, J. de Jesus Rangel-Magdaleno, J. M. Ramirez-Cortes, H. Peregrina-Barreto and R. Morales-Caporal, "Empirical Mode Decomposition Analysis for Broken-Bar Detection on Squirrel Cage Induction Motors," *IEEE Trans. Instrum. Meas.*, vol. 64, no. 5, pp. 1118-1128, May 2015.

[10] B. Wang, J. Wang, A. Griffo and B. Sen, "Stator Turn Fault Detection by Second Harmonic in Instantaneous Power for a Triple-Redundant Fault-Tolerant PM Drive," *IEEE Trans. Ind. Electron.*, vol. 65, no. 9, pp. 7279-7289, Sept. 2018.

[11] A. Sapena-Baño et al., "Harmonic Order Tracking Analysis: A Novel Method for Fault Diagnosis in Induction Machines," *IEEE Trans. Energy Convers.*, vol. 30, no. 3, pp. 833-841, Sept. 2015.

[12] D. Cao, W. Hu, J. B. Zhao, Q. Huang, Z. Chen, F. Blaabjerg, "A multi-agent deep reinforcement learning based voltage regulation using coordinated PV inverters," *IEEE Trans. Power Syst.*, vol. 35, no. 5, pp. 4120-4123, Sept. 2020.

[13] D. Cao, J. B. Zhao, W. Hu, et al.., "Attention enabled multi-agent DRL for decentralized volt-VAR control of active distribution system using PV inverters and SVCs" *IEEE Trans. Sustain. Energy*, early access.

[14] D. Cao, W. Hu, J. Zhao, et al. "Reinforcement learning and its applications in modern power and energy systems: a review." *Journal of Modern Power Systems and Clean Energy*, vol. 8, no. 6, pp. 1029-1042, 2020.

[15] J. Chen et al., "Novel Data-Driven Approach Based on Capsule Network for Intelligent Multi-Fault Detection in Electric Motors," *IEEE Trans. Energy Convers.*, early access.

[16] J. Chen, W. Hu, D. Cao, et al., "An Imbalance fault detection algorithm for variable-speed wind turbines: A deep learning approach," *Energies*, vol. 12, no. 14, pp. 2764-2778, Jul. 2019.

[17] F. Wang, R. Liu, Q. Hu and X. Chen, "Cascade Convolutional Neural Network with Progressive Optimization for Motor Fault Diagnosis Under Non-stationary Conditions," *IEEE Trans. Ind. Inform.*, early access.

[18] R. Liu, F. Wang, B. Yang and S. J. Qin, "Multiscale Kernel Based Residual Convolutional Neural Network for Motor Fault Diagnosis Under Nonstationary Conditions," *IEEE Trans. Ind. Inform*., vol. 16, no. 6, pp. 3797-3806, June 2020.

[19] S. Shao, R. Yan, Y. Lu, P. Wang and R. X. Gao, "DCNN-Based Multi-Signal Induction Motor Fault Diagnosis," *IEEE Trans. Instrum. Meas.*, vol. 69, no. 6, pp. 2658-2669, June 2020.

[20] J. Lei, C. Liu, D. Jiang, "Fault diagnosis of wind turbine based on Long Short-term memory networks," *Renewable Energy*, vol. 133, pp. 422-432, Oct. 2018.

[21] Z. An, S. Li, J. Wang, X. Jiang, "A novel bearing intelligent fault diagnosis framework under time-varying working conditions using recurrent neural network," *ISA Transactions*, vol. 100, pp. 155-170, Nov. 2019.

[22] L. Wang, Z. Zhang, J. Xu and R. Liu, "Wind Turbine Blade Breakage Monitoring with Deep Autoencoders," *IEEE Trans. Smart Grid*, vol. 9, no. 4, pp. 2824-2833, July 2018.

[23] F. Cheng, J. Wang, L. Qu and W. Qiao, "Rotor-Current-Based Fault Diagnosis for DFIG Wind Turbine Drivetrain Gearboxes Using Frequency Analysis and a Deep Classifier," *IEEE Trans. Ind. Appl*., vol. 54, no. 2, pp. 1062-1071, March-April 2018.

[24] M. Ma, C. Sun and X. Chen, "Deep Coupling Autoencoder for Fault Diagnosis with Multimodal Sensory Data," *IEEE Trans. Ind. Informa*., vol. 14, no. 3, pp. 1137-1145, March 2018.

[25] P. Liang, C. Deng, J. Wu, G. Li, Z. Yang and Y. Wang, "Intelligent Fault Diagnosis via Semisupervised Generative Adversarial Nets and Wavelet Transform," *IEEE Trans. Instrum. Meas.*, vol. 69, no. 7, pp. 4659-4671, July 2020.

[26] S. Shao, S. McAleer, R. Yan and P. Baldi, "Highly Accurate Machine Fault Diagnosis Using Deep Transfer Learning," *IEEE Trans. Ind. Inform.*, vol. 15, no. 4, pp. 2446-2455, April 2019.

[27] N. Lai, M. Kan, C. Han, X. Song and S. Shan, "Learning to Learn Adaptive Classifier-Predictor for Few-Shot Learning," *IEEE Trans. Neural Netw. Learn. Syst.*, doi: 10.1109/TNNLS.2020.3011526.

[28] A. Zhang, S. Li, Y. Cui, W. Yang, R. Dong and J. Hu, "Limited Data Rolling Bearing Fault Diagnosis with Few-Shot Learning," *IEEE Access*, vol. 7, pp. 110895-110904, 2019.

[29] S. Zhang, F. Ye, B. Wang, T. G. Habetler, "Few-Shot Bearing Anomaly Detection Based on Model-Agnostic Meta-Learning," arXiv preprint arXiv:2007.12851

[30] J. Wu, Z. Zhao, C. Sun, et al., "Few-shot transfer learning for intelligent fault diagnosis of machine," *Measurement*, vol. 166, Dec. 2020.

[31] K. He, X. Zhang, S. Ren, J. Sun, "Deep Residual Learning for Image Recognition," *arXiv preprint arXiv: 1512.03385v1*, 2015.

[32] M. Patacchiola, J. Turner, E. J. Crowley, et al., "Deep Kernel Transfer in Gaussian Processes for Few-shot Learning," *arXiv preprint arXiv:1910.05199v3*, 2020.

[33] W. Chen, Y. Liu, Z. Kira, et al., "A closer look at few-shot classification," *arXiv preprint arXiv: 1904.04232*, 2019.

[34] D. Silver, A. Huang, C. J. Maddison, et al., "Mastering the game of Go with deep neural networks and tree search," *Nature*, vol. 529, no. 5787, pp. 484-489, Jan. 2016.

[35] Case Western Reserve University Bearing Data Center. [Online]. Available: http://csegroups.case.edu/bearingdatacenter/pages/welcome-case-western-reserve-university-bearing-data-center-website.

[36] C. Finn, P. Abbeel, S. Levine, "Model-Agnostic Meta-Learning for Fast Adaptation of Deep Networks," *arXiv preprint arXiv: 1703.03400v3*, 2017.

[37] J. Snell, K. Swersky, R. S. Zemel, "Prototypical Networks for Few-shot Learning," *arXiv preprint arXiv: 1703.05175v2*, 2017.

[38] W. Chen, Y. Liu, Z. Kira, "A Closer Look at Few-shot Classification," *arXiv preprint arXiv: 1904.04232v2*, 2020.

[39] W. Zhang, G. Peng, C. Li, Y. Chen, and Z. Zhang, "A new deep learning model for fault diagnosis with good anti-noise and domain adaptation ability on raw vibration signals," *Sensors*, vol. 17, no. 2, pp. 425–446, 2017.

[40] S. Sabour, N. Frosst, and G. E. Hinton. "Dynamic routing between capsules," *arXiv preprint arXiv: 1710.09829*, 2017.

[41] Hatem A. Fayed, Amir F. Atiya. "Speed up grid-search for parameter selection of support vector machines," Applied Soft Computing, vol. 80, pp. 202-210, July 2019.

[42] I. M. Abdulbaqi, A. T. Humod, O. K. Alazzawi, "Application of FEM to Provide the Required Database for MCSA Based On-line Fault Detection System on 3-phase Induction Motor using ANSYS Maxwell2D," *Advances in Natural & Applied Sciences*, vol. 10, no. 16 pp. 43-54, Nov. 2016.

Jianjun Chen received the B.S. from Chongqing University of Technology, Chongqing, China, in 2018. He is currently working toward the Ph.D. degree in control science and engineering at the University of Electronic Science and Technology of China (UESTC). His research interests include mechanical fault diagnosis and deep learning.

Weihao Hu (S'06–M'13–SM'15) received the B.Eng. and M.Sc. degrees from Xi'an Jiaotong University, Xi'an, China, in 2004 and 2007, respectively, both in electrical engineering, and Ph. D. degree from Aalborg University, Denmark, in 2012.

He is currently a Full Professor and the Director of Institute of Smart Power and Energy Systems (ISPES) at the University of Electronics Science and Technology of China (UESTC). He was an Associate Professor at the Department of Energy Technology, Aalborg University, Denmark and the Vice Program Leader of Wind Power System Research Program at the same department. His research interests include artificial intelligence in modern power systems and renewable power generation. He has led/participated in more than 15 national and international research projects and he has more than 170 publications in his technical field.

He is an Associate Editor for IET Renewable Power Generation, a Guest Editor-in-Chief for Journal of Modern Power Systems and Clean Energy Special Issue on Applications of Artificial Intelligence in Modern Power Systems, a Guest Editor-in-Chief for Transactions of China Electrical

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TEC.2021.3070024, IEEE Transactions on Energy Conversion

Technology Special Issue on Planning and operation of multiple renewable energy complementary power generation systems, and a Guest Editor for the IEEE TRANSACTIONS ON POWER SYSTEM Special Section on Enabling very high penetration renewable energy integration into future power systems. He was serving as the Technical Program Chair (TPC) for IEEE Innovative Smart Grid Technologies (ISGT) Asia 2019 and is serving as the Conference Chair for the Asia Energy and Electrical Engineering Symposium (AEEES 2020). He is currently serving as Chair for IEEE Chengdu Section PELS Chapter. He is a Fellow of the Institution of Engineering and Technology, London, U.K. and an IEEE Senior Member.

Di Cao received the B.S. from North China University of Water Resources and Electric Power, Zhengzhou, China, in 2014. He is currently working toward the Ph.D. degree in control science and engineering at the University of Electronic Science and Technology of China. His research interest includes optimization of distribution network.

Man Zhang (M'20) and received the B.S. and M.S. degrees from Northwestern Polytechnical University, Xi'an, China, in 2012 and 2015, respectively. She prepared the Ph.D. degree at Génie Electrique et électronique de Paris, Gif-Sur-Yvette, France, and received the Ph.D. degree from University Paris-XI, France, in 2019, all in electrical engineering. She is currently a Lecture with the School of Mechanical and Electrical Engineering, University of Electronic Science and Technology of China. Her research interests include numerical and analytical analysis of multiphysic problems, vibration, and noise reduction of SRM

Qi Huang (S'99, M'03, SM'09) was born in Guizhou province in the People's Republic of China. He received his BS degree in Electrical Engineering from Fuzhou University in 1996, MS degree from Tsinghua University in 1999, and Ph.D. degree from Arizona State University in 2003. He is currently a professor at UESTC, the Executive Dean of School of Energy Science and Engineering, UESTC, and the director of Sichuan State Provincial Lab of Power System Wide-area Measurement and Control. He is a member of IEEE since 1999. His current research and academic interests include power system instrumentation, power system monitoring and control, and power system high performance computing.

Zhe Chen (M'95-SM'98-F'19) received the B.Eng. and M.Sc. degrees from Northeast China Institute of Electric Power Engineering, Jilin, China, and the Ph.D. degree from University of Durham, Durham, U.K.

He is a Full Professor with the Department of Energy Technology, Aalborg University, Denmark. He is the Leader of Wind Power System Research Program in the Department of Energy Technology, Aalborg University and the Danish Principle Investigator for Wind Energy of Sino-Danish Centre for Education and Research. His research areas include power systems, power electronics and electric machines; and his main current research interests are wind energy and modern power systems. He has led many research projects and has more than 400 publications in his technical field.

Dr. Chen is an Editor of the IEEE TRANSACTIONS ON POWER SYSTEMS, an Associate Editor of the IEEE TRANSACTIONS ON POWER ELECTRONICS, a Fellow of the Institution of Engineering and Technology, London, U.K., a Chartered Engineer in the U.K., and a Fellow of the IEEE.

Frede Blaabjerg (S'86–M'88–SM'97–F'03) was with ABB-Scandia, Randers, Denmark, from 1987 to 1988. From 1988 to 1992, he got the PhD degree in Electrical Engineering at Aalborg University in 1995. He became an Assistant Professor in 1992, an Associate Professor in 1996, and a Full Professor of power electronics and drives in 1998. From 2017 he became a Villum Investigator. He is honoris causa at University Politehnica Timisoara (UPT), Romania and Tallinn Technical University (TTU) in Estonia.

His current research interests include power electronics and its applications such as in wind turbines, PV systems, reliability, harmonics and adjustable speed drives. He has published more than 600 journal papers in the fields of power electronics and its applications. He is the co-author of four monographs and editor of ten books in power electronics and its applications.

He has received 32 IEEE Prize Paper Awards, the IEEE PELS Distinguished Service Award in 2009, the EPE-PEMC Council Award in 2010, the IEEE William E. Newell Power Electronics Award 2014, the Villum Kann Rasmussen Research Award 2014, the Global Energy Prize in 2019 and the 2020 IEEE Edison Medal. He was the Editor-in-Chief of the IEEE TRANSACTIONS ON POWER ELECTRONICS from 2006 to 2012. He has been Distinguished Lecturer for the IEEE Power Electronics Society from 2005 to 2007 and for the IEEE Industry Applications Society from 2010 to 2011 as well as 2017 to 2018. In 2019-2020 he serves as President of IEEE Power Electronics Society. He is Vice-President of the Danish Academy of Technical Sciences too. He is nominated in 2014-2019 by Thomson Reuters to be between the most 250 cited researchers in Engineering in the world.