Article:
Yang, D, Pang, Y, Zhou, B et al. (1 more author) (2019) Fault Diagnosis for Energy Internet Using Correlation Processing-Based Convolutional Neural Networks. IEEE Transactions on Systems, Man, and Cybernetics: Systems, 49 (8). pp. 1739-1748. ISSN 2168-2216

https://doi.org/10.1109/TSMC.2019.2919940

Fault Diagnosis for Energy Internet Using Correlation Processing Based Convolutional Neural Networks

***Abstract*—Fault feature extraction based on prior knowledge and raw data is increasingly becoming more challenging in energy internet fault diagnosis due to complicated network topology and coupling disturbances introduced into the systems. Deep learning methods that have emerged in recent years, such as the convolutional neural networks (CNNs), have shown a number of advantages and great potentials in the field of feature extraction and image recognition. However, CNNs does not work well in fault diagnosis for industrial systems, due to the totally different data representations between images used in recognition and signals obtained from industrial processes. This paper tackles this problem by introducing a novel and generic fault diagnosis method for complicated system namely the Spearman rank correlation based CNNs (SR-CNN). By imposing the Spearman rank correlation image layer on the typical CNNs, the multiple time series signals measured by the phasor measurement units (PMUs) is converted to appropriate data images, which are then fed to the CNNs. With the aid of this novel design, different fault features can be comprehensively extracted while the fault can be identified more quickly and precisely than other conventional approaches. To validate the efficacy of the proposed approach, an IEEE defined power gird with multi new energy resources are used as the test platform. The experimental results confirm the effectiveness and superiority of the proposed method in energy internet fault diagnosis over conventional methods.**

***Index Terms*—Fault diagnosis, convolutional neural network, multiple data processing, spearman rank correlation.**

I. Introduction

FAULT diagnosis is a critical and non-trivial task in complex system network applications such as energy internet to improve the reliability and reduce the loss caused by a large number of potential incidents [1], [2]. The fault diagnosis algorithms and methods introduced so far can be categorized as follows [3]

- mechanism model based analysis
- analysis based on expertise and prior knowledge
- wide area monitoring based signal processing

The mechanism model is often built on the basis of the physical principles and around some fault scenarios, such as the voltage sag or the power flow. The model is thus established to monitor the consistency between the measured outputs and the model outputs [4]–[8]. However, these methods become increasingly difficult to detect faults occurring in the energy internet, since the network complexity increasing significantly due to the introduction of distributed renewable generations and a variety of new loads such as electric vehicles. The mechanism model for analysis or control thus often becomes too complicated to build [9], [10]. In addition, single model can only be used to detect a single type of fault, which limits its applications for the energy internet.

Fault analysis based on expertise and prior knowledge can handle some situations where the acquired information is incomplete [11]–[13]. Yet, the complexity of energy internet in terms of the size and coupling is growing exponentially in recent years due to the roll-out of a large number of distributed generations and loads. These coupled characters among subsystems and components are under-researched and not explicit at this stage. That means this kind of methods can hardly keep in pace with the exponentially growing network complexity, and its applicability can be even worse.

Signal processing method utilizes widely measured signals using specific equipment such as phasor measurement units and the Supervisory Control And Data Acquisition (SCADA). Instead of exploring the mechanism or the model of the fault, this kind of method can analyze the features of the acquired signals and associate them with the fault state. Numerous algorithms have been proposed so far, such as the wavelet analysis, spectral analysis, and model analysis, etc [14]–[19]. However, these traditional methods still request some prior knowledge and demand significant amount of on-line computational resources. These methods are becoming increasingly difficult to handle fault diagnosis for energy internet applications where the sheer system complexity presents the most challenging bottleneck.

Deep learning method is originated from the artificial neural networks (ANNs) field and it offers a greater potential for complex signal processing. In artificial neural networks, multiple hidden layers are contained to recognize features and patterns [20]. While particularly in the image processing and pattern recognition field, the convolutional neural network is a typical application of deep learning method.

CNNs based methods for fault diagnosis are proposed in recent years, there are many researches of applying CNNs to one-dimensional (1D-CNN) signal. For example, the literature [21] proposes a novel intelligent diagnosis method to identify different health conditions of wind turbine gearbox. A novel method is proposed to automatically select the impulse responses from the vibration signals for early fault detection in [22]. The literature [23] detects the fault of rolling element bearing using the vibration signals. A feature learning model is proposed for condition monitoring based on CNNs in [24]. A fault identification and classification method for gearboxes based on its vibration signals is presented in [25].

However, other researchers believe that these one-dimensional based deep learning methods cannot work well in complex system network application, since methods like CNNs are proposed for image recognition, and there are great differences between the two-dimensional images (2D

images) and the raw signals [26]. In other words, the biggest challenge for CNNs based fault diagnosis is how to overcome the information presentation where traditional CNNs is mainly developed for images, while most fault diagnosis applications in the energy internet are based on time series signals.

To address this challenge, the literature [27] develops a novel diagnosis framework called dislocated time series CNNs to dislocate the input signals and establish the corresponding image representation. The literature [28] presents a conversion method converting signals into 2D representations and then extract the fault features in a manufacturing system. A novel computational intelligence-based electrocardiogram signal classification methodology using a deep learning machine is developed in [29]. The literature [30] presents a method for fault diagnosis and fault selection based on CNNs using the current signals in a power transmission system. The literature [31] introduces an architecture that automatically learns a robust set of feature representations from raw spatio-temporal tomography sensor data. The literature [32] presents a novel deep convolutional neural network cascading architecture for performing localization and detecting defects in insulators. A time-frequency gray scale images are acquired by applying the continuous wavelet transform (CWT) and used for earth fault detection in [33].

These above reported methods and applications only use a single type of signals or images for feature extraction. However, in practical applications of the energy internet, many variables of system networks are measured and recorded, each of them contains some partial fault features. That means the fault features are distributed in multiple signals irregularly, and this kind of complicated non-linear relationship is hard to analyze. That requires a comprehensive analysis using all available signal types with novel data mining techniques.

This paper proposes a novel fault diagnosis method, namely the SR-CNN. The rationale is that the variations in the Spearman rank correlation can reflect the feature of different faults. Therefore, the massive time series signals measured in energy internet can be converted to 2D images to allow for feature extraction and image recognition in CNNs. The proposed method not only overcomes the challenge when the CNNs is applied to the fault diagnosis in energy internet applications, but also increases the interpretability of the recognized results. Further, the proposed method presents a great potential to handle fault diagnosis for systems and networks with increasing complexity in terms of multiple signals and strong coupling of elements.

The primary contributions of this paper can be summarized as follows.

- A new and generic method namely SR-CNN is proposed to analyze faults in complex system network using multiple signals based on deep learning method. It is especially suitable for energy internet because of its structural characteristic and preferable performance.
- The proposed method offers a number of merits:
  - more accurate fault diagnosis eliminating the impact of distributed generations, direction of power flow and fault impedance;
  - conversion of multiple signals to a series of 2D images for CNNs based feature recognition;
  - faster diagnosis of multiple fault types than conventional approaches;
  - requires no priori knowledge or empirical analysis of faults.
- The method can be broadly used for signal processing with a wide range of system signals, which contains complicated and coupled characteristics.

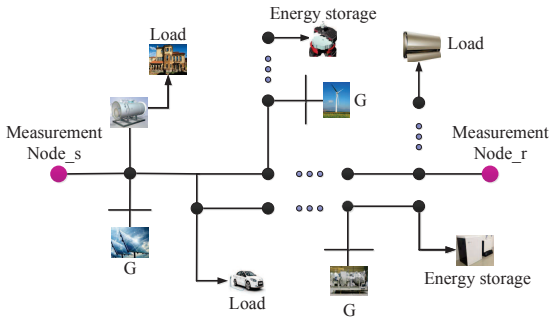


Fig. 1. The schematic of PMUs based wide area monitoring system in energy internet

The rest of this paper is organized as follows. In Section II, the measurement configuration in application scenario is presented together with the correlation theory. In Section III, the proposed SR-CNN framework and its technical details are given. Section IV presents the case studies of the proposed methods with experimental results. Section V concludes the paper.

II. Data Collection and Spearman Correlation

A. Measurement Configuration in Application Scenario

Wide area measurement system (WAMS) using Phasor Measurement Units (PMUs) has been widely deployed worldwide in recent years. WAMS can measure and transmit multiple signals according to the GPS synchronous clock. Several kinds of variables can be measured by PMUs, including node voltage, branch current, phase angle, active and reactive power, and so on. These signals are gathered at every sampling instant and transmitted to the master station. Therefore, the synchronized data can be grouped and stored. Fig.1 is a schematic of the wide area monitoring system using PMUs.

In this schematic, there are two types of measurement nodes in the regional network, namely the sending nodes and the receiving nodes. Even though the direction of the power flow can be changed due to the introduction of distributed generators, it does not affect the performance of the proposed method, because the model established by this method is purely based on the PMUs signals. Therefore, the whole energy internet consists of several measurement structures like this. The optimal PMUs placement has been discussed [34]–[37]. This paper assumes that the locations of the measurement nodes are appropriate such that the collected signals contain sufficient information of faults in the energy internet.

B. Spearman Rank Correlation

Dependency analysis is an important step for data processing. The product-moment correlation (often referred to Pearson product-moment correlation) and rank correlation (often also referred to the Spearman's Rank-Order Correlation) are two crucial dependency analysis approaches. The Pearson product-moment assumes normal distribution or approximately normal distribution of the observations, which is not suitable for variables measured by PMUs in energy internet. While, the rank correlation does not need such assumptions. For each two vectors $X_{n\times1}$ and $Y_{n\times1}$, the Spearman rank correlation can be calculated as

$$\begin{aligned}\rho &= \frac{cov(x,y)}{\sigma_x\sigma_y} \\ &= \frac{E[(x-\mu_x)(y-\mu_y)]}{\sigma_x\sigma_y} \\ &= \frac{\sum_{i=1}^{n} x_i y_i - n\bar{x}\bar{y}}{\sqrt{(\sum_{i=1}^{n} x^2{}_i - \bar{x}^2)(\sum_{i=1}^{n} y^2{}_i - \bar{y}^2)}}\end{aligned} \tag{1}$$

where, x or y are the rank vectors of the raw vectors X and Y, which are measured synchronized electric signals from the PMUs at the sending and the receiving node, respectively. μ, σ are the mean value and variance, respectively. Spearman correlation assesses monotonic relationships between two rank variables.

Since the rank is a consecutive positive integer, $\bar{x}, \bar{y}, \sum_{i=1}^{n} x^2, \sum_{i=1}^{n} y^2, \sum_{i=1}^{n} x_i y_i$ can be calculated as [38]

$$\begin{aligned}\bar{x} = \bar{y} &= \frac{1}{n}(1+2+\cdots+n) = \frac{n+1}{n} \\ \sum_{i=1}^{n} x^2 &= \sum_{i=1}^{n} y^2 = 1^2+2^2+\cdots+n^2 \\ &= \frac{n(n+1)(2n+1)}{6} \\ \sum_{i=1}^{n} x_i y_i &= \frac{1}{2}\sum_{i=1}^{n}[x^2{}_i + y^2{}_i - (x_i-y_i)^2] \\ &= \frac{n(n+1)(2n+1)}{6} - \frac{1}{2}\sum_{i=1}^{n} d^2{}_i\end{aligned} \tag{2}$$

where,$d_i = x_i - y_i$. Therefore, (1) can be rewritten as

$$\rho = 1 - \frac{6\sum {d_i}^2}{n(n^2-1)} \tag{3}$$

Intuitively, the correlation between two variables will be higher when the observations have a similar rank (i.e. relative position label of the observations within the variable: 1st, 2nd, 3rd, etc.), and lower when observations have a dissimilar (or fully opposite for a correlation of 1) rank between the two variables. Thus, the Spearman coefficient is appropriate for both continuous and discrete ordinal variables [39].

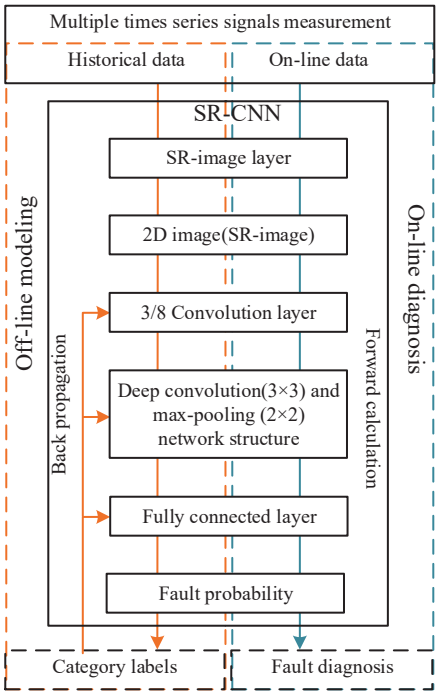


Fig. 2. The overall flow chart of fault diagnosis

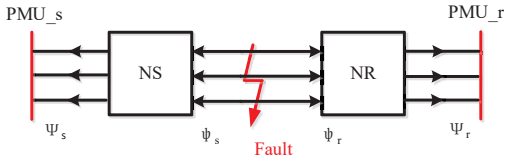


Fig. 3. The topological structure of an equivalent fault network

III. Proposed Fault Diagnosis Method

This section proposes a novel method for signal processing with applications to the energy internet wide area monitoring. By imposing a Spearman rank correlation image layer on the conventional CNNs, the multiple PMUs signals can be converted to more appropriate 2D images to improve the fault diagnosis performance. The modeling process of the proposed SR-CNN is an off-line supervised learning and the trained model analyze the fault using on-line signals. The overall flow chart of fault diagnosis is shown in Fig.2.

A. Relationship Between Fault Features and Spearman Rank Correlation

In order to illustrate the relationship between the fault features and the Spearman rank correlation qualitatively, and to demonstrate the feasibility of fault diagnosis based on the SR-CNN, an equivalent fault network is presented. When the fault occurs between two measurement nodes, the distributed parameters can be viewed as two nonlinear functions from the fault position. The equivalent fault network is illustrated in Fig.3.

Supposing a fault occurs at the position indicated in the equivalent network, the influence caused by the fault spreads to both the sending node and the receiving node. Thus, the

regularity they follow as be formulated as two nonlinear functions, denoted as $\mathcal{NS}$ and $\mathcal{NR}$, which obviously depends on the type of the fault as well as the type of the signal. According to the superposition principle, the signal from the sending node and the receiving node can be represented as

$$\begin{aligned} \Psi_s &= \psi_s \oplus \mathcal{NS}_{\psi}(fault) \\ \Psi_r &= \psi_r \oplus \mathcal{NR}_{\psi}(fault) \end{aligned} \tag{4}$$

where, Ψ_s and Ψ_r are the after-fault signal from both the sending and the receiving node, respectively. ψ_s and ψ_r are the pre-fault signal from both the sending and the receiving node, respectively. Symbol $\oplus$ is the vector addition operation that adds each element of vector 1 to the corresponding element of vector 2. Therefore, the spearman rank correlation of Ψ_s and Ψ_r can be depicted as

$$\begin{aligned} \rho' &= 1 - \frac{6\sum({x^2}_{\Psi_s,i} - {y^2}_{\Psi_r,i})}{n(n^2-1)} \\ &= 1 - \frac{6\sum D_i^2}{n(n^2-1)} \\ where, D_i^2 &= [(x_{\psi_s,i} + x_{\mathcal{NS},i}) - (y_{\psi_r,i} + y_{\mathcal{NR},i})]^2 \end{aligned} \tag{5}$$

The vector dimension is fixed and $\sum D_i^2$ is the key factor to reveal the relationship between the spearman rank correlation and the fault features. According to the associative law and the commutative law, $\sum D_i^2$ can be depicted as

$$\begin{aligned} \sum D_i^2 &= \sum[(x_{\psi_s,i} - y_{\psi_r,i}) + (x_{\mathcal{NS},i} - y_{\mathcal{NR},i})]^2 \\ &= \sum {d^2}_{\psi,i} + \sum[2d_{\Psi,i}d_{\psi,i} + {d^2}_{\Psi,i}] \\ where, d_{\psi,i} &= x_{\psi_s,i} - y_{\psi_r,i} \\ d_{\Psi,i} &= x_{\mathcal{NS},i} - y_{\mathcal{NR},i} \end{aligned} \tag{6}$$

Therefore, the variation of spearman rank correlation between pre-fault signals and after-fault signals can be formulated as

$$\begin{aligned} \Delta\rho &= \rho - \rho' \\ &= (1 - \frac{6\sum {d^2}_{\psi,i}}{n(n^2-1)}) - (1 - \frac{6\sum D_i^2}{n(n^2-1)}) \\ &= \frac{6\sum(2d_{\Psi,i}d_{\psi,i} + {d^2}_{\Psi,i})}{n(n^2-1)} \end{aligned} \tag{7}$$

It is clear that $\Delta\rho$ is related to $d_{\Psi,i}$, which are related to the nonlinear functions $\mathcal{NS}$ and $\mathcal{NR}$. Thus, $\Delta\rho$ retains reveals the fault features.

Traditionally, a mechanism model is established to analyze this change quantitatively. However, the increasingly complicated topology, distributed generations and massive loads as well as a large quantity of coupling disturbances subsequently introduced into the system significantly increase the difficulty of modeling and diagnosis. In addition, these mechanism methods request large amounts of on-line computation resources and one model can only detect one fault type. The applicability of mechanism models is becoming very limited.

Machine learning is generally considered to be capable of handling massive data with great potentials to unveil the coupling relationships among a large amount of variables. It can use off-line data to model the patterns of variations between the Spearman rank correlation and fault features, then perform fault diagnosis using online data based on the trained model. The difficulties associated with traditional approaches can thus be overcome.

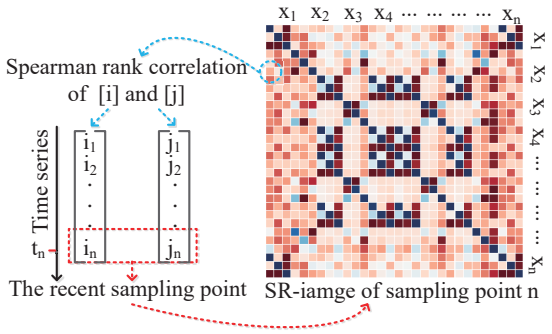


Fig. 4. The Spearman rank correlation image (SR-image) built by SR-image layer

B. Spearman Rank Correlation Image Layer

Data preprocessing plays an important role in data-driven methods. In this proposed method, for the signal vector $i_{n\times1}$ and $j_{n\times1}$, which are two kinds of system signal consisted of n sample points, the correlation coefficient ρ_{ij} can be calculated based on (3) using the rank of these vectors. Suppose that the serial number i and j are the position coordinates, and each two signal vectors of the energy internet can obtain an element ρ. Therefore, the Spearman rank correlation matrix can be established as

$$SR - matrix = \begin{bmatrix} \rho_{11} & \cdots & \rho_{1n} \\ \vdots & \ddots & \vdots \\ \rho_{n1} & \cdots & \rho_{nn} \end{bmatrix} \tag{8}$$

However, each pixel of the 2D image is an integer between 0 and 255. Thus, the scaling transformation needs to be employed to satisfy this requirement. For each element ρ_{ij} of the Spearman rank correlation matrix, its corresponding pixel can be calculated as

$$P_{ij} = \mathcal{R}(\frac{\rho_{ij} - \rho_{min}}{\rho_{max} - \rho_{min}} \times 255) \tag{9}$$

where, P_{ij} is the pixel of the SR-image and $\mathcal{R}$ is the rounding function. ρ_{min} and ρ_{max} are the minimum and the maximum of Spearman rank correlation, respectively.

Therefore, the Spearman rank correlation image layer (SR-image layer) converts the multiple raw signals to an appropriate 2D images at each sample point, as illustrated in Fig.4.

Compared with the data images established based on the time series data using the traditional method, the advantages of SR-image can be summarized as follows.

- Within the local receptive field, the correlation of pixels is enhanced and the distribution is more regular. These kind

of characters will significantly enhance the performance and the efficiency of fault recognition by CNNs.
- The locations of fault features in the images are fixed, hence the recognition is not affected by the fault occurring time.
- The number of fault features is likely increased, hence helps to train the convolution kernels more easily and quickly. The increased number of fault features also eases the situation where insufficient fault data is available.
- The SR-images are also able to contain a large amount of signals as the correlations analysis does not limit the data in terms of size and dimension, in contrast to the traditional methods.

C. Spearman Rank Correlation Based Convolutional Neural Architecture

The specific SR-CNN architecture established in this paper is graphically illustrated in Fig.5. According to the PMUs data collected, the SR-image layer processes these data by correlation calculations and image reconstruction. Its output generates suitable images for the initial convolutional layer. Let $z = 3/8$ defining that there are 8 convolution kernels with a 3×3 initial convolutional layer, and $x = y$ represent the size of an SR-image. Each kernel filters the input image and produces a feature map. The second convolutional layer, which is defined as $3/16$, filters the feature map produced by the previous layer. Hence, the fault features are extracted from a local field to the advanced field. The process of convolution can be formulated as

$$C(x,y) = \sum_{-\infty}^{\infty}\sum_{-\infty}^{\infty} \mathcal{K}(n,n) * \mathcal{F}(x-n, y-n) \tag{10}$$

where, $\mathcal{K}$ is the convolution kernel, and $\mathcal{F}$ is the object matrix for convolution.

The volume of feature maps increases rapidly after twice convolution and some redundant information are produced. Therefore, the down-sampling operation is proceeded by a 2×2 max-pooling layer. Then, the third convolutional layer defined as $3/32$ is used to extract the global feature. After another max-pooling, the feature map is flattened to a large vector. The fully connected layer maps the complex nonlinear relationship by its hidden neurons. The activation function used in this framework is defined as

$$f(x) = \begin{cases} x, & x >= 0 \\ \alpha x, & x < 0 \end{cases} \tag{11}$$

where, α is a small constant. By using this activation function, also often named as the leaky ReLU, the information from the negative axis can be partially reserved and thus the vanishing gradient can be relieved.

The output vector of a fully connected layer is used as the input of a softmax layer and the final output vector is the probability of diagnosis for different fault types. Apparently, a fault that occurs in the energy internet presents the highest probability value. The optimization objective of the proposed architecture is the convolution kernels and the connection weights. Besides, the loss function is defined by the cross entropy as

$$\begin{aligned} \mathcal{J}(\theta) = &-\frac{1}{m}\sum_{i=1}^{m} y^{(i)} \log(h_\theta(x^{(i)})) \\ &+ (1-y^{(i)}) \log(1 - h_\theta(x^{(i)})) \end{aligned} \tag{12}$$

where, $h_\theta(x^{(i)}))$ is the hypothesis function, and it is defined as

$$h_\theta(x^{(i)})) = \frac{1}{1 + e^{-\theta^T x^{(i)}}} \tag{13}$$

Therefore, the optimization is proceeded by taking the derivative of $\mathcal{J}(\theta)$ and then modifying the kernels and weights.

In summary, in the proposed architecture, appropriate 2D images are constructed using the PMUs data and the fault features can be extracted and recognized using the deep neural network. The SR-CNN architecture not only takes full advantage of massive and multiple dimensional signals obtained by the WAMS, but also promote the wide-spread applications of CNNs in energy internet environment.

TABLE I
FAULT TYPES IN THE EXPERIMENT MODEL

Numbers	FaultType
0	Normal state
1	Single-phase ground fault
2	Two-phase short circuit fault
3	Two-phase ground fault
4	Three-phase short circuit fault
5	Three-phase ground fault

IV. EXPERIMENTS

A. Simulation Model and Parameterization

The standard IEEE 14-bus power system is established using the PSCAD/EMTDC to verify effectiveness and superiority of the proposed method. The network structure and configuration of the simulation model is shown in Fig.6. In this model, different fault types occur at four positions respectively and the fault types are listed in Table 1. In addition, the frequency of this standard model is 60Hz, so the sampling frequency is 6 kHz which is achievable because the PMUs can sample ten thousand points per second. According to the Nyquist's sampling law, this sampling frequency is greater than the duple system frequency, so the sample data can express the system state. In addition, the generator power and the load power in this experiment are also shown in Fig.6.

B. Fault Data Image Construction

The fault type 1, 2, and 4 take place at location 1, respectively. The fault occurring time is set as 0s. When the fault occurs, a massive number of signal samples are available

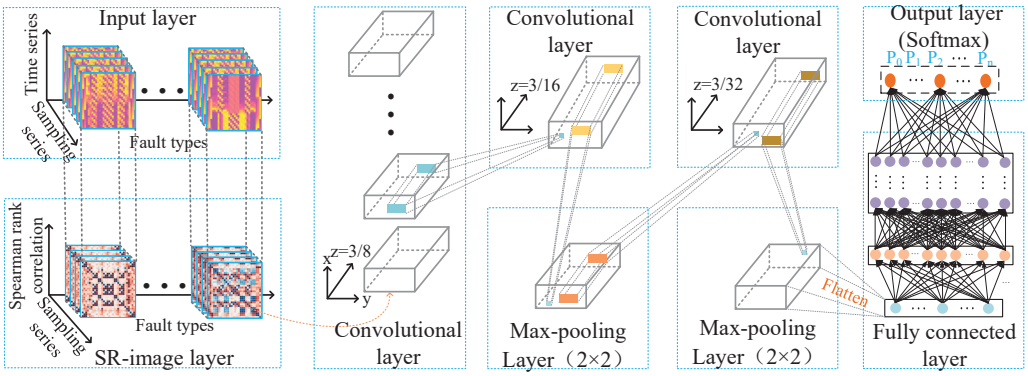


Fig. 5. The proposed Spearman rank correlation CNNs (SR-CNN) architecture

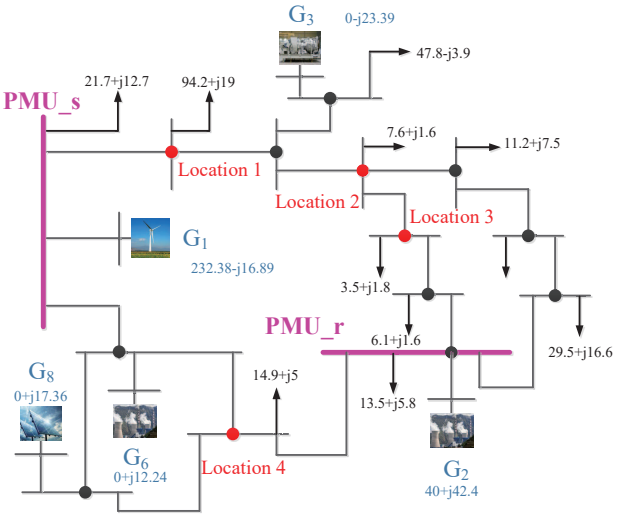


Fig. 6. Simulation model built in PSCAD/EMTDC

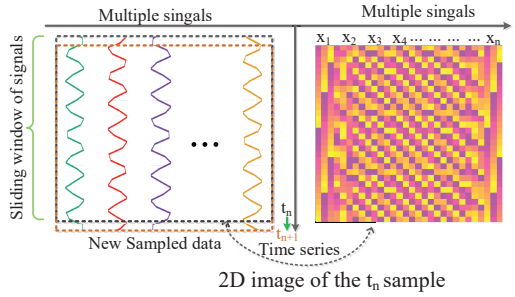


Fig. 7. The fault data image established by the traditional methods

covering each sampling instant, thus can be used to detect the fault.

Several methods have been proposed to convert time-series data to data images [27], [28]. Although these methods perform well with vibration signals in industrial application, they are hard to deal with multiple signals. According to these papers, the fault data image can be established and illustrated in Fig.7. The gray image is converted to the heat map in this paper in order to reveal the data feature more clearly.

As shown in Fig.7, the three phase electrical quantities (active power, reactive power, phase angle, branch current, node voltage) from both the sending node and the receiving node constitute the x-axis, while a window of time series data of the electrical variables constitute the y-axis. Therefore, a fault data image is equivalent resembles a moving window of data, and as a new data sample moves in, the very first data is then removed.

Though this kind of data images meet the fundamental input requirement for CNNs, they are far from the optimal representation of the raw data for effective and efficient fault diagnosis. For example, the magnitudes of these variables on the x-axis are significantly different and there is few correlation between the adjacent element on the x-axis. These characters lead to weak correlations of information at local receptive field in the 2D image, thus impact the performance of CNNs for fault recognition.

Further, for time series data, the location of fault features can vary significantly if this kind of data images are used. Fig.8 illustrates the fault data images established by the traditional methods at 0.05s, 0.10s and 0.15s of different faults. There exhibits only one obvious fault feature in these images and they can be viewed as feature lines rather than feature maps, which do not reveal two-dimensional structure. This implies that the multi-level and nonlinear extraction of fault features using the convolution kernels is not appropriate enough. The data distribution is also skewed due to the differences between multiple variables. Besides, fault type 1 does not reveal any obvious feature at 0.05s in the basic data images. It is also difficult to distinguish the features between fault type 1 and fault type 4 at 0.10s and 0.15s. The accuracy of fault recognition will be affected beacuse CNNs can only handle the location variation to some extent [40].

To overcome the aforementioned problems with this kind of images, the fault data represented in the SR-image layer is illustrated in Fig.9. Apparently, the SR-images are proper 2D images, which reveal strong correlations in the local receptive field. These images not only contain several evident features for fault diagnosis, but also the locations of these features in the images are fixed. In addition, after the fault occurs, the distinction of different faults is also evident even in a very short time, so the diagnosis can be made rapidly, hence offering great potential to improve the system reliability. Fig.10 illustrates the SR-image of all six states at 0.15s.

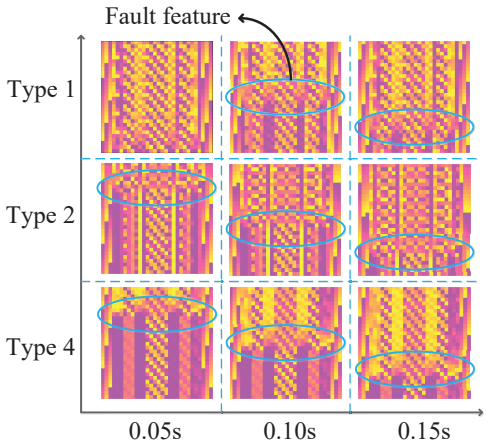


Fig. 8. Images of different fault types established by the traditional methods

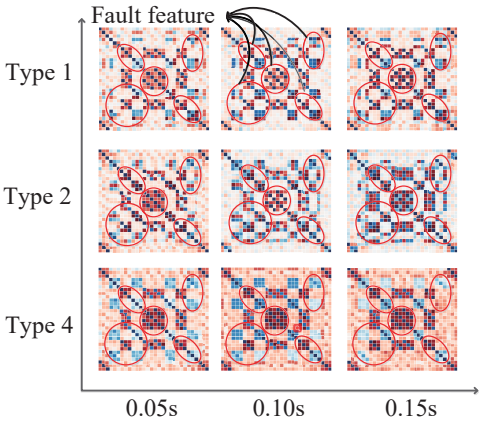


Fig. 9. The SR-image of different fault types at different sampling point

C. Validation of SR-CNN

In order to verify the performance of SR-CNN comprehensively, three levels validation are used: efficiency based on insufficient data, accuracy based on massive data, and timeliness of fault diagnosis. Further, the universal training method, namely batch gradient descent (BGD) is used to obtain the optimal results. Table 2 lists three sets of data and the basic parameters for verification. The fault will be diagnosed when its diagnostic probability is maximal in the softmax vector, which can be formulated as

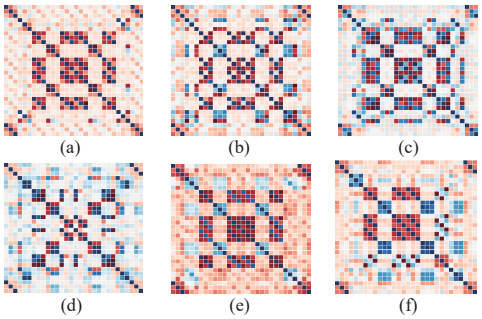


Fig. 10. The SR-image of different fault types at 0.15s. (a) Type 0. (b) Type 1. (c) Type 2. (d) Type 3. (e) Type 4. (f) Type 5.

TABLE II
THREE SETS OF ESSENTIAL DATA FOR VERIFICATION

Set	Location	Types	Data Size
1	1	1-5	1500
2	1,2,3	1-5	9800
3	4	1-5	1500

$$Faulttype = Max(softmax) = Max(P_0, \cdots, P_5) \quad (14)$$

Firstly, the efficiency of fault diagnosis based on the insufficient data is tested. In this scenario, only the first set of data in Table 2 is used for training and testing. In order to evaluate the performances of the proposed SR-CNN, some other methods are selected to compare performance in this case. They are 1D signal based CNNs (1D-CNN) [21], a conversion method based CNNs (CMCNN) [28], and Google's VGG11 fed with the SR-image. The k-fold cross validation method is used for the comparison, two conditions $k = 5$ and $k = 10$ are taken into account. Moreover, thirty times experiments are carried out with fine tuning the model parameters, such as the number of hidden layers, the activation function and the learning rate. In each step, the data set has randomly shuffled to ensure the fairness. The average results of the accuracy, the number iterations and the proportion of convergence are shown in Table 3.

The SR-CNN peaks at about 67th epoch and its average accuracy is about 11 percent higher than the 1D-CNN. Since this new designed deep learning model is more suitable for the recognition of time series data such as vibration signal, and not appropriate for the image based analysis for mutiple signals in the energy internet application. The result also shows that the image recognition is more suitable for multiple signals analysis than the data filtering in the complex system application. Furthermore, the average accuracy of SR-CNN is about 7 percent higher than the CMCNN and 8 percent higher than the VGG11. Meanwhile, the SR-CNN also shows the significant performance on the ability of convergence compared with other methods.

Generally speaking, the convolution kernels and the weights of CNNs need to be trained based on enormous data to extract specific fault features. It is clear that modeling process can be shortened if the fault features are easy to recognize. Further, the increasing number of useful fault features can fully take the advantage of the multiple convolutional layers, so that all of their kernels are used to distinguish the different features even though the data is insufficient. This implies that the SR-CNN can make more accurate diagnosis even the available data is insufficient. This characteristic can be particularly useful when the historic fault data is no longer sufficient.

Secondly, massive data are used to verify the accuracy of the fault diagnosis of SR-CNN. In this test, the modeling is based on the extension of the previous step. Therefore, the epoch of training can be shortened. The second set of data is added for training and the results are illustrated in

TABLE III
COMPARISON RESULT BASED ON THE K-FOLD CROSS VALIDATION

Methods	Average Accuracy		Iterations	Proportion
	$k = 5$	$k = 10$		
SR-CNN	0.9275	0.9474	67 ± 3	93.3%
CMCNN	0.8513	0.8772	65 ± 5	73.3%
1D-CNN	0.8146	0.8391	71 ± 5	86.7%
VGG11	0.8336	0.8625	65 ± 2	63.3%

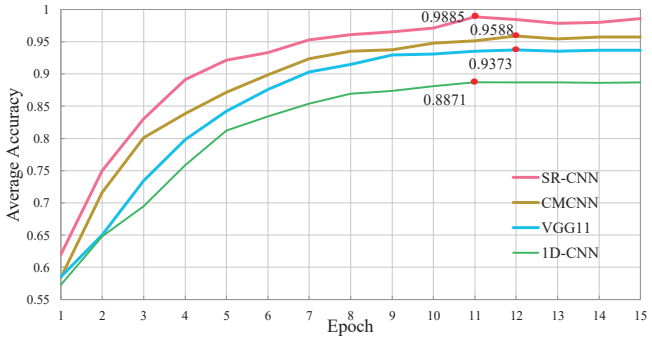


Fig. 11. Comparison of accuracy based on the massive data

Fig.11. It is clear that, the accuracy of SR-CNN is better than the other methods: about 10 percent higher than the 1D-CNN and at least 3 percent higher than other methods. That demonstrates its learning and generalization capacity. That means the proposed network has proper topology and can take full advantage of massive data, extract the feature of multiple variables and make a very precise fault diagnosis for the energy internet.

Thirdly, the final set of data is used to verify the timeliness of fault diagnosis using the model developed in the previous step. Fig.12 illustrates the probability of diagnosis for different faults type in time series. It is clear that, the diagnostic probability of the normal state decreases with time when the fault occurs at 0s in the system. While the probability of the fault, which has occurred, increases gradually until reaching close to one. More specifically, for fault type 3 and 4, the probability is almost 0.9 at 0.05s and while for type 1 and 2 faults at 0.10s. For fault type 5, the probability increases at a slightly slower pace because its fault feature is close to the type 4. However, the fault is confirmed only when its probability reaches the highest. Thus, the fault type 5 can also be detected precisely at 0.15s by SR-CNN. Further, Fig.13 compares the average timeliness of fault diagnosis and other mehtods. It confirms that SR-CNN is able to produce rapid and accurate fault diagnosis in the energy internet applications.

According to the average results shown in Fig.13, more than 60 percent accurate diagnosis can be achieved by SR-CNN about 0.10s after the fault occurs and it increases to more than 90 percent when the time is 0.15s. This efficient result can be illustrated by two aspects: on the one hand, the proposed SR-image layer can explore and express the fault feature at the very beginning, earlier than the traditional methods; on the

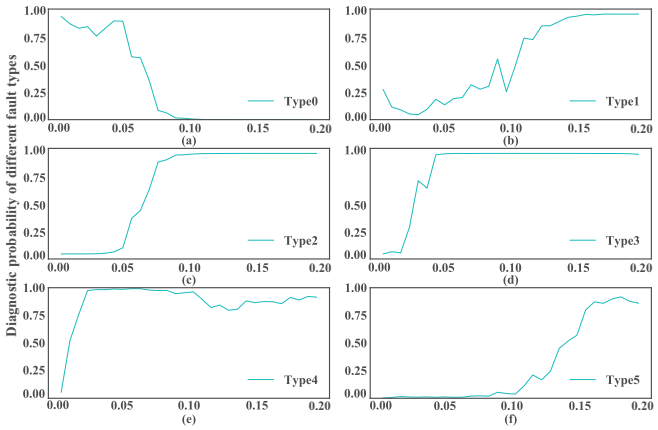
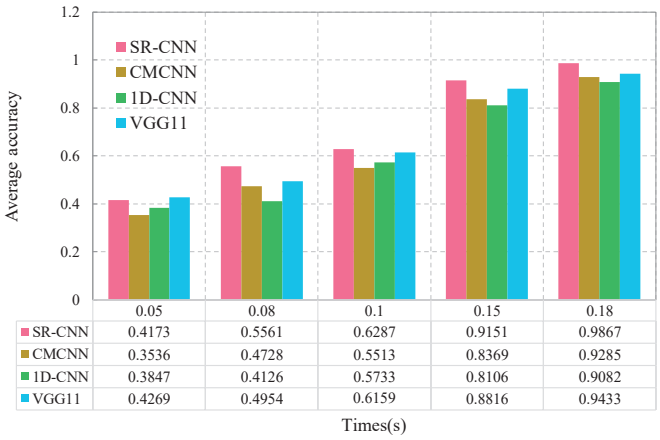


Fig. 12. The sequential probability of different fault types at 0.15s. (a) Type 0. (b) Type 1. (c) Type 2. (d) Type 3. (e) Type 4. (f) Type 5



	0.05	0.08	0.1	0.15	0.18
SR-CNN	0.4173	0.5561	0.6287	0.9151	0.9867
CMCNN	0.3536	0.4728	0.5513	0.8369	0.9285
1D-CNN	0.3847	0.4126	0.5733	0.8106	0.9082
VGG11	0.4269	0.4954	0.6159	0.8816	0.9433

Fig. 13. Comparison of the timeliness for rapid diagnosis

other hand, the location of fault features in the established 2D image is settled, so that the convolution process, which has not any attention mechanism, can reach its ideal effect. That confirms the ability of SR-CNN for early detection of the fault in energy internet applications.

As for the influence of measurement uncertainty like packet loss, which is caused by the restriction of signal strength or transmission equipment, can also be relieved by the SR-CNN to some extent. It is because that the proposed SR-image layer calculates each pixel using two different vectors, and this kind of vectors contains a time series sampled point. That means the SR-image represents the relative variation tendency of multiple signals in a period of time and a handful of loss or redundancy has little influence on the image features. In general, the proposed method can make an accurate diagnosis even using the incomplete data, but the accuracy rate will decline along with the increase of measurement uncertainty.

V. CONCLUSION

In this paper, a novel and generic fault diagnosis method, namely SR-CNN is proposed for complicated system network, especially in energy internet applications. Benefiting from the imposition of the SR-image layer, the multiple variables measured by PMUs can be converted to generate 2D images

suitable for fast recognition by CNNs, benefiting from the increasing number, fixed position and obvious identity of fault features. The SR-CNN not only can retain the superiority of CNNs on image recognition, but also take full advantages of the massive data made available acquired by PMUs in the energy internet. The experimental results confirm that the SR-CNN is able to offer accurate and fast fault diagnosis even with insufficient data. Further, the diagnosis accuracy and timeliness are also investigated, and the results confirm the precision and rapidity of the SR-CNN based fault diagnosis in the energy internet. The proposed method can serve as the reference for fault diagnosis in the complex systems, so the network structure can be adjusted appropriately for the specific application scenarios. Moreover, there are still something need to improve of the proposed approach. For example, since its modeling is a supervised learning process, the classification of training data has a great influence on the accuracy of analysis. That also means the approach cannot analyze the fault type, which has not contained in the training data or has no appropriate category labels. That may be the direction of further research.

REFERENCES

[1] M. Kezunovic, "Smart fault location for smart grids," *IEEE Transactions on Smart Grid*, vol. 2, no. 1, pp. 11–22, March 2011.

[2] D. Yang, X. Gao, Z. Ma, E. Cui, and H. Zhang, "Novel voltage sag protection topology of contactors for uninterrupted switching capability," *IEEE Transactions on Industry Applications*, vol. 54, no. 4, pp. 3170–3178, July 2018.

[3] Z. Gao, C. Cecati, and S. X. Ding, "A survey of fault diagnosis and fault-tolerant techniquespart ii: Fault diagnosis with knowledge-based and hybrid/active approaches," *IEEE Transactions on Industrial Electronics*, vol. 62, no. 6, pp. 3768–3774, June 2015.

[4] L. Xu and H. E. Tseng, "Robust model-based fault detection for a roll stability control system," *IEEE Transactions on Control Systems Technology*, vol. 15, no. 3, pp. 519–528, May 2007.

[5] F. Caccavale, A. Marino, G. Muscio, and F. Pierri, "Discrete-time framework for fault diagnosis in robotic manipulators," *IEEE Transactions on Control Systems Technology*, vol. 21, no. 5, pp. 1858–1873, Sep. 2013.

[6] S. Kang, Y. Ahn, Y. Kang, and S. Nam, "A fault location algorithm based on circuit analysis for untransposed parallel transmission lines," *IEEE Transactions on Power Delivery*, vol. 24, no. 4, pp. 1850–1856, Oct 2009.

[7] R. N. Mahanty and P. B. D. Gupta, "Application of rbf neural network to fault classification and location in transmission lines," *IEE Proceedings - Generation, Transmission and Distribution*, vol. 151, no. 2, pp. 201–212, March 2004.

[8] X. Cong, M. P. Fanti, A. M. Mangini, and Z. Li, "Decentralized diagnosis by petri nets and integer linear programming," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 48, no. 10, pp. 1689–1700, Oct 2018.

[9] C. Dou, D. Yue, and J. M. Guerrero, "Multiagent system-based event-triggered hybrid controls for high-security hybrid energy generation systems," *IEEE Transactions on Industrial Informatics*, vol. 13, no. 2, pp. 584–594, April 2017.

[10] S. Weng, D. Yue, C. Dou, J. Shi, and C. Huang, "Distributed event-triggered cooperative control for frequency and voltage stability and power sharing in isolated inverter-based microgrid," *IEEE Transactions on Cybernetics*, pp. 1–13, 2018.

[11] Z. Yi and A. H. Etemadi, "Fault detection for photovoltaic systems based on multi-resolution signal decomposition and fuzzy inference systems," *IEEE Transactions on Smart Grid*, vol. 8, no. 3, pp. 1274–1283, May 2017.

[12] H. Peng, J. Wang, J. Ming, P. Shi, M. J. Prez-Jimnez, W. Yu, and C. Tao, "Fault diagnosis of power systems using intuitionistic fuzzy spiking neural p systems," *IEEE Transactions on Smart Grid*, vol. 9, no. 5, pp. 4777–4784, Sep. 2018.

[13] X. Wang, S. D. J. McArthur, S. M. Strachan, J. D. Kirkwood, and B. Paisley, "A data analytic approach to automatic fault diagnosis and prognosis for distribution automation," *IEEE Transactions on Smart Grid*, vol. 9, no. 6, pp. 6265–6273, Nov 2018.

[14] L. Hong and J. S. Dhupia, "A time domain approach to diagnose gearbox fault based on measured vibration signals," *Journal of Sound and Vibration*, vol. 333, no. 7, pp. 2164–2180, Mar 2014.

[15] Y. Gritli, L. Zarri, C. Rossi, F. Filippetti, G. Capolino, and D. Casadei, "Advanced diagnosis of electrical faults in wound-rotor induction machines," *IEEE Transactions on Industrial Electronics*, vol. 60, no. 9, pp. 4012–4024, Sep. 2013.

[16] V. Climente-Alarcon, J. A. Antonino-Daviu, M. Riera-Guasp, and M. Vlcek, "Induction motor diagnosis by advanced notch fir filters and the wignerville distribution," *IEEE Transactions on Industrial Electronics*, vol. 61, no. 8, pp. 4217–4227, Aug 2014.

[17] Z. Feng and M. J. Zuo, "Fault diagnosis of planetary gearboxes via torsional vibration signal analysis," *Mechanical Systems & Signal Processing*, vol. 36, no. 2, pp. 401–421, 2013.

[18] G. M. Joksimovi, J. Riger, T. M. Wolbank, N. Peri, and M. Vaak, "Stator-current spectrum signature of healthy cage rotor induction machines," *IEEE Transactions on Industrial Electronics*, vol. 60, no. 9, pp. 4025–4033, Sep. 2013.

[19] S. Lu, Q. He, T. Yuan, and F. Kong, "Online fault diagnosis of motor bearing via stochastic-resonance-based adaptive filter in an embedded system," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 47, no. 7, pp. 1111–1122, July 2017.

[20] J. Schmidhuber, "Deep learning in neural networks: An overview," *Neural Networks*, vol. 61, pp. 85–117, Jan 2015.

[21] G. Jiang, H. He, J. Yan, and P. Xie, "Multiscale convolutional neural networks for fault diagnosis of wind turbine gearbox," *IEEE Transactions on Industrial Electronics*, vol. 66, no. 4, pp. 3196–3207, April 2019.

[22] B. Luo, H. Wang, H. Liu, B. Li, and F. Peng, "Early fault detection of machine tools based on deep learning and dynamic identification," *IEEE Transactions on Industrial Electronics*, vol. 66, no. 1, pp. 509–518, Jan 2019.

[23] F. Example, "Deep learning enabled fault diagnosis using time-frequency image analysis of rolling element bearings," *Shock & Vibration*, vol. 2017, pp. 1–17, 2017.

[24] O. Janssens, V. Slavkovikj, B. Vervisch, K. Stockman, M. Loccufier, S. Verstockt, R. V. D. Walle, and S. V. Hoecke, "Convolutional neural network based fault detection for rotating machinery," *Journal of Sound & Vibration*, vol. 377, pp. 331–345, 2016.

[25] Z. Chen, C. Li, and R. Snchez, "Gearbox fault identification and classification with convolutional neural networks," *Shock and Vibration*, vol. 2015, pp. 1–10, 10 2015.

[26] H. Jeong, S. Park, S. Woo, and S. Lee, "Rotating machinery diagnostics using deep learning on orbit plot images," *Procedia Manufacturing*, vol. 5, pp. 1107–1118, 2016.

[27] R. Liu, G. Meng, B. Yang, C. Sun, and X. Chen, "Dislocated time series convolutional neural architecture: An intelligent fault diagnosis approach for electric machine," *IEEE Transactions on Industrial Informatics*, vol. 13, no. 3, pp. 1310–1320, June 2017.

[28] L. Wen, X. Li, L. Gao, and Y. Zhang, "A new convolutional neural network-based data-driven fault diagnosis method," *IEEE Transactions on Industrial Electronics*, vol. 65, no. 7, pp. 5990–5998, July 2018.

[29] B. Pourbabaee, M. J. Roshtkhari, and K. Khorasani, "Deep convolutional neural networks and learning ecg features for screening paroxysmal atrial fibrillation patients," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 48, no. 12, pp. 2095–2104, Dec 2018.

[30] D. Wei, Q. Gong, W. Lai, B. Wang, D. Liu, H. Qiao, and G. Lin, "Research on internal and external fault diagnosis and fault-selection of transmission line based on convolutional neural network," *Proceedings of the Csee*, vol. 36, pp. 21–28, 09 2016.

[31] O. Costilla-Reyes, P. Scully, and K. B. Ozanyan, "Deep neural networks for learning spatio-temporal features from tomography sensors," *IEEE Transactions on Industrial Electronics*, vol. 65, no. 1, pp. 645–653, Jan 2018.

[32] X. Tao, D. Zhang, Z. Wang, X. Liu, H. Zhang, and D. Xu, "Detection of power line insulator defects using aerial images analyzed with convolutional neural networks," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, pp. 1–13, 2018.

[33] M. Guo, X. Zeng, D. Chen, and N. Yang, "Deep-learning-based earth fault detection using continuous wavelet transform and convolutional neural network in resonant grounding distribution systems," *IEEE Sensors Journal*, vol. 18, no. 3, pp. 1291–1300, Feb 2018.

[34] E. Khorram and M. T. Jelodar, "Pmu placement considering various arrangements of lines connections at complex buses," *International*

Journal of Electrical Power & Energy Systems, vol. 94, pp. 97–103, 2018.

[35] S. Zhu, W. Lei, S. Mousavian, and J. H. Roh, "An optimal joint placement of pmus and flow measurements for ensuring power system observability under n-2 transmission contingencies," *International Journal of Electrical Power & Energy Systems*, vol. 95, pp. 254–265, 2017.

[36] M. T. Jelodar and A. S. Fini, "Probabilistic pmu placement considering topological change in high voltage substations," *International Journal of Electrical Power & Energy Systems*, vol. 82, pp. 303–313, 2016.

[37] Y. Zhao, Y. Peng, A. Qian, and T. Lv, "Optimal pmu placement considering topology constraints," *International Journal of Electrical Power & Energy Systems*, vol. 73, pp. 240–248, 2015.

[38] J. L. Myers, A. D. Well, and J. Lorch, R. F., *Research Design and Statistical Analysis (2nd ed.)*. Lawrence Erlbaum, 1995.

[39] A. Lehman, N. O'Rourke, L. Hatcher, and E. Stepanski, *Jmp For Basic Univariate And Multivariate Statistics: A Step-by-step Guide*, 2013.

[40] O. Janssens, V. Slavkovikj, B. Vervisch, K. Stockman, M. Loccufier, S. Verstockt, R. V. D. Walle, and S. V. Hoecke, "Convolutional neural network based fault detection for rotating machinery," *Journal of Sound & Vibration*, vol. 377, pp. 331–345, 2016.