

Article

## **Convolutional Neural Network-Based Stator Current Data-Driven Incipient Stator Fault Diagnosis of Inverter-Fed Induction Motor**

Maciej Skowron<sup>®</sup>, Teresa Orlowska-Kowalska \*<sup>®</sup>, Marcin Wolkiewicz<sup>®</sup> and Czeslaw T. Kowalski<sup>®</sup>

Department of Electrical Machines, Drives and Measurements, Wroclaw University of Science and Technology, 50-370 Wroclaw, Poland; maciej.skowron@pwr.edu.pl (M.S.); marcin.wolkiewicz@pwr.edu.pl (M.W.); czeslaw.t.kowalski@pwr.edu.pl (C.T.K.)

\* Correspondence: teresa.orlowska-kowalska@pwr.edu.pl; Tel.: +48-71-320-2640

Received: 22 February 2020; Accepted: 18 March 2020; Published: 20 March 2020



**Abstract:** In this paper, the idea of using a convolutional neural network (CNN) for the detection and classification of induction motor stator winding faults is presented. The diagnosis inference of the stator inter-turn short-circuits is based on raw stator current data. It offers the possibility of using the diagnostic signal direct processing, which could replace well known analytical methods. Tests were carried out for various levels of stator failures. In order to assess the sensitivity of the applied CNN-based detector to motor operating conditions, the tests were carried out for variable load torques and for different values of supply voltage frequency. Experimental tests were conducted on a specially designed setup with the 3 kW induction motor of special construction, which allowed for the physical modelling of inter-turn short-circuits in each of the three phases of the machine. The on-line tests prove the possibility of using CNN in the real-time diagnostic system with the high accuracy of incipient stator winding fault detection and classification. The impact of the developed CNN structure and training method parameters on the fault diagnosis accuracy has also been tested.

**Keywords:** fault diagnosis; induction motor drive; inter-turn short circuits; convolutional neural network; deep learning

### 1. Introduction

Nowadays, drive systems based on AC (alternative current) electric motors play a key role in industrial applications. It is estimated that these machines account for approximately 29% of global and 69% of industrial electricity consumption. During the operation of electric motors in industrial drive systems, various types of defects may occur, preventing further operation of the machine. According to EPRI's statistics, the most common defects of electrical machines include stator (37%), rotor (10%), and rolling bearing (41%) damages. The conducted static tests show that, with the increase of the rated power of electric motors, the role of mechanical damages in favor of electrical damages decreases [1]. This fact causes that many scientific centers are focusing on developing diagnostic methods, enabling a partial reduction of the number of electrical damages or detecting the state of the machine (incipient fault), in which it can still be repaired.

The basis for the operation of diagnostic systems is a thorough knowledge of the changes occurring in the machine as a result of damage. Observation of changes occurring in AC machines is carried out using signals available for measurement on the tested object. The most commonly used are currents [2,3], voltages [4,5], vibrations [6,7] as well as flux [8,9] and temperature [10,11]. The idea of analytical methods for assessing the technical condition of the machine is based on the extraction of damage symptoms in measured diagnostic signals.



The use of the analytical approach in the fault detection processes of electrical windings of the AC motors is connected with the necessity to adjust the parameters of diagnostic signal measurement to the method of fault symptom extraction. Currently used analytical methods of damage detection are based on the observation of signal changes in the domain of time [12], frequency [13] or time-frequency [14]. Among them, the following methods deserve special attention: fast Fourier transform (FFT) [15,16], extended park's vector analysis (EPVA) [17], short-time Fourier transform (STFT) [14,18], Hilbert–Huang transform (HHT) [19,20] as well as wavelet analysis in the form of continuous wavelet transform (CWT) [21,22] or discrete wavelet transform (DWT) [23]. Of the methods presented, the most common is the use of FFT analysis. The signal spectrum analysis provides both an effective assessment of damage symptoms and does not constitute a significant burden on computing systems. However, like most analytical methods, FFT requires a relatively long measurement time, which, especially in the case of inter-turn short circuits of a stator winding, is an undoubted disadvantage of this method, due to the dynamics of these faults. Moreover, it is in many cases not possible to ensure signal stationarity over the required measurement period.

Recognition of changes in diagnostic signals and their assessment in systems based on analytical methods is usually performed by a human being, which increases the time for detecting faults. Moreover, such systems are often not fully automated, and the number of errors is often linked to the experience of the human expert.

Currently, due to the dynamic development of artificial intelligence methods, diagnostic systems based on analytical methods are gradually replaced by neuronal damage detectors. The undoubted advantage of these systems is the limited role of the human expert in the process of assessing the object's technical condition, while ensuring the high efficiency of fault detection. The basis for the operation of diagnostic systems based on neural networks (NNs), applied also for induction motor (IM) fault detection and classification, are analytical methods. Therefore, the input information for such systems is the result of extracting symptoms from chosen diagnostic signals using analytical methods, for example: FFT [24,25], WT [26], HHT [27], etc.

Multilayer neural networks are the most popular structures used in diagnostic techniques, also in the case of induction and synchronous motor fault detection. The use of a multilayer perceptron (MLP) was discussed in detail in [28–30]. The MLP network acts as a training data approximator, mapping the learning data with the lowest possible level of a cost function. In addition to data approximation capabilities, a particularly important task of NN-based diagnostic systems is damage classification. The main representative of classifiers used in fault detection processes of IMs is the Kohonen self-organizing network (SOM) [31–33]. The advantage of SOM is its simple structure and the fact that there is no need to ensure multiple learning samples. Due to their structure, SOM based diagnostic applications are used, in particular to recognize the category of damages [34].

Artificial intelligence methods are increasingly associated with the issue of deep learning. The basis of deep learning structures are classic neural networks. However, deep neural networks (DNN) are characterized by features different from those of classical shallow networks, which results from the abandonment of the universal approximation rule. The structures used so far, with a maximum of two hidden layers, have allowed one to achieve satisfactory results [29,31]. Nevertheless, increasing the structure made it possible to obtain NN features that were before unachievable in classical shallow structures.

For a long time, DNNs were used mainly in information systems (e.g., image processing, speech recognition). Over the past three years, there has been a growing interest in DNN-based electrical machine damage detection systems noticed. For the most part, they are associated with the analysis of mechanical damage to the IM [35,36] or mechanical system components [37,38]. A small number of works related to electrical damage of IM mainly concern rotor damages [24,27]. Most DNN-based systems use vibration measurements [39–41], less frequently stator currents [42–44] and voltages [45]. This fact results from clear changes occurring in the diagnostic signal, due to a mechanical damage and the resulting simplicity of the signal analysis.

Among the used deep learning network structures, the most commonly used ones are convolutional neural networks (CNNs) [46–49] and autoencoders [50,51]. CNNs used in diagnostic applications are characterized by a much higher level of effectiveness compared to MLP and RBF networks, which are presented in [25,49]. In diagnostic processes, CNNs can act as a damage classification system [49], as well as provide information about the degree of damage [24]. In [51], the possibility of using CNNs as members of an autoencoder structure was presented. By comparing structures based on MLP, CNN and LSTM (long short-term memory), the authors [51] showed that the use of an appropriate structure of autoencoders significantly affects the effectiveness of the diagnostic system. An important aspect in the use of DNN in diagnostic processes is the appropriate adjustment of the measured signal to the structure and properties of the network. The input vector of DNNs can result from the signal analysis [24,25] and the directly provided diagnostic signal [35,46]. Due to the principle of operation of deep learning structures, in most cases the measured signal is converted into a 2D [25,48,52] or 3D [51] matrix. The impact of the input matrix size on the damage detection efficiency is presented, among others, in [48].

This article presents the possibility of using the CNNs in the detection process of inter-turns short circuits in IM stator windings, especially incipient faults. The developed application is characterized by the direct processing of a raw diagnostic signal. The authors presented the possibility of extracting damage symptoms directly from the stator phase current signals, omitting well known analytical pre-processing methods, such as the FFT, WT, and other higher-order transforms [53].

The idea of the CNN-based diagnostic system that performs the task of the fault detecting and assessing the degree of damage of the IM stator windings is described. Unlike the methods presented in [35,46] based on an easy to analyze vibration signal for the detection of mechanical damages, the developed method consists of the direct processing of the phase current signals through a DNN, which is a novelty in the detection systems of stator faults. It should be emphasized that the development of a diagnostic system based on direct analysis of the stator current is associated with additional difficulties, due to the fact that the increase in the number of shorted turns causes a similar effect as an increase in the load torque, i.e., an increase in stator current amplitude. However, the diagnostic system should detect the characteristics of the damage, i.e., distinguish between the effects of short circuits and the impact of the drive operating conditions on the diagnostic signal being analyzed. Nevertheless, the article shows the high efficiency of the developed technique for detecting incipient damage to the stator winding of the converter-fed induction machine.

The article is divided into five main sections. The introduction is followed by the second section presenting the theoretical basis of the developed CNN structure. In the third section, the discussion on diagnostic signal processing method and the training process, as well as the parameters of the developed CNN structure applied for the incipient stator fault detection of the IM is presented. The laboratory set-up is also presented in the section. In the next section, the results of experimental verification are presented, including the on-line detection of inter-turn short circuit faults in the one phase and in three stator phases of the motor operating in different conditions. In the following fifth part, the authors focused on the impact of CNN training process parameters and the proposed network structure on the effectiveness of the developed diagnostic system. The conclusions and observations resulting from the performed research are presented in the last section.

#### 2. General Description of Convolutional Neural Network Used in the Research

#### 2.1. Structure of the Convolutional Network

The basic function of CNN is to extract the features of higher orders from the analyzed signal using convolutional operations. These networks do not have a preconceived architecture, parameter selection methods or rules regarding the number of convolutional layers.

A CNN structure should be seen as the determination of features, which progress with each additional convolutional layer. In the presented CNN application in the diagnostic process, the

first layer can be understood as the filter of basic features, e.g., maximum or minimum values. The subsequent execution of the convolutional operation allows us to detect higher order features, i.e., the distances between minimum and maximum values. Therefore, the network structure will depend on the type of information provided, as well as the function performed by CNN.

To detect complex features, structures composed of multiple sets of layers are used. The ability to detect features is associated with the process of acquiring generalization skills by the network. Due to the extensive structure of CNN, the methods which avoid over-matching are becoming very important. The structure and method of CNN training will be described below on the example of the network used in this paper, for the diagnosis of IM stator winding incipient damages.

The following Figure 1 presents an example of basic structure of the CNN consisting of a few sets of convolutional layers and one set of layers responsible for determining class membership. The task of the CNNs developed in the application described in this paper was to distinguish the degree of damage to the IM stator windings on the basis of information derived directly from the raw phase current signal (details will be given in the next section).



Figure 1. The basic structure of the convolutional neural network.

The principle of the CNN's individual layers was discussed in detail, among others, in [54–61]. The application of each of the presented layers enabled CNN to acquire the characteristic features of the stator current data, which resulted in an increase in the efficiency of the IM faults neuronal detector. The convolutional layer [54,55] performs the function of a feature detector using the convolutional operation of combining two data sets. This convolutional operation is performed for multidimensional input arrays. In the application of the CNN for stator fault detection presented in this paper, the convolutional layer acts as a filter searching for fault symptoms in the phase current waveform.

To accelerate the training process and also to increase the stability of the NN training, the batch normalization layer was used [56]. This layer normalizes the output of the previous layer, subtracting the average value of the batch elements and dividing this received value by their standard deviation. The impact of using this normalization method is thoroughly discussed in [56]. As in the case of classic NN structures, the activation function plays a crucial role in DNN structures. The most commonly used activation function is the rectified linear unit (ReLU) [54,57]. It is mainly used as a complement to convolutional layers and allows us to capture interactions and representing nonlinearities.

Convolutional layers provide a very large amount of information (symptoms) observed in the input matrix. In many cases, the use of so many object features is pointless. Pooling layers are used in a similar way [54,58]. Their task is to choose only the information whose contribution to each cell (window) is the largest. For this purpose, methods that search for maximum or average values from cell elements are most often used. The advantage of pooling layers is that they reduce the spatial size of data representation, thus preventing overmatching.

The significant number of CNN parameters causes some difficulties in giving the neurons special characteristics, from the perspective of generalization property. The dropout layer [54,59,60] is used to avoid a situation in which a single neuron strongly depends on the state of the others. This technique allows one to teach each of the neurons a different useful feature of the analyzed input data. Due to the rejection of a number of neural connections in each iteration, the training process is accelerated, and overmatching is prevented.

The last layer in the CNNs is usually the fully connected layer, which enables the assessment of the share of individual classes that are the result of the network operation [54,55]. Fully connected layers applied in CNNs mostly cooperate with softmax layers [61]. Their task is to determine the probability of the input batch elements belonging to one of the categories. The softmax layers' response takes the form of a vector of probabilities and is used by the classification layer to determine the belonging of the CNN input matrix to one of the specified categories. For this purpose, the cross entropy of losses is calculated.

### 2.2. Training of the Convolutional Network

The CNN training process, in most cases, is carried out according to the algorithm of stochastic gradient descent (SGD). The SGD training method allows one to determine the unloaded gradient estimation, using, for this purpose, the average of sample gradients drawn from the mini data packet. The most important parameter of the SGD algorithm is the learning speed. This value is most often chosen in trial and error performed by analyzing learning curves. Unfortunately, this approach does not allow for the optimization of the learning process, because if the value of this parameter is too high, it will cause the rapid oscillation of the learning curve while the value results in extended training time will be too low. An alternative to the fairly slow SGD algorithm is its extension with the momentum parameter (SGDM). Unlike the SGD method, the SGDM algorithm has the largest step size when multiple consecutive gradients point in the same direction [62]. The training process, according to the SGDM algorithm, begins with determining the initial value of the learning rate  $\eta$  and the momentum parameter  $\alpha$ . Then, for the sample from the training mini packet { $x_1, \ldots, x_k$ }, the gradient p is estimated according to the relationship:

$$p = \frac{1}{k} \nabla_{\mathbf{w}} \sum_{i} L(f(\mathbf{x}_{i}, \mathbf{w}), \mathbf{y}_{i}),$$
(1)

where:  $\mathbf{x}_i$ —a randomly selected minibatch element size m,  $L(f(\mathbf{x}_i, \mathbf{w}), \mathbf{y}_i)$ —calculated loss function for the *i*-th sample.

For estimated gradient value *p*, momentum **v** and parameter **w** are updated in accordance with the following relationships:

$$\mathbf{v}(k) = \alpha \cdot \mathbf{v}(k-1) - \eta(k) \cdot \mathbf{p}(k),$$
  

$$\mathbf{w}(k) = \mathbf{w}(k-1) + \mathbf{v}(k),$$
(2)

where:  $\alpha$ —hyperparameter determining how quickly the contributions of previous gradients disappear exponentially.

The above algorithm is repeated until the stop condition is met, which is most often the number of training periods or the value of the loss function. The SGDM method, thanks to the use of data packets to approximate the gradient estimation, enables network training during experimental verification, using the data from current measurements. This is an advantage of this method, especially in terms of implementing CNNs in electric machine fault detection systems.

# 3. Description of the Experimental Implementation of the CNN-Based Detector of Inter-Turn Short Circuits

### 3.1. Short Description of the Laboratory Set-Up

The experimental verification of the proposed CNN-based fault detector was carried out on a specially designed setup with IM of 3 kW, presented in Figure 2. The construction of the used IM allowed for the physical modelling of inter-turn short circuits in each of the three phases of IM in a range of 0–5 stator turns (parameters and the nominal data of the IM, as well as the connection of the stator winding for inter-turn short-circuit modelling are shown in the Appendices A and B). The tested motor was fed from a frequency converter in the range  $f_s = 10-50$  Hz, operating in an open speed loop under scalar control  $u_s/f_s = \text{const.}$  The tests were carried out for various values of the load torque in the

range  $T_L = (0-1)T_N$ . The loading machine was fed from an industrial frequency converter, operating in a torque mode, with 10 kHz modulation frequency.



**Figure 2.** Experimental set-up: (**a**) general scheme of the set-up; (**b**) components of the real experimental drive system.

During the preparation of the CNN training data, the information about the actual values of stator currents was used. The measured diagnostic signals were provided by the data acquisition measurement card (DAQ) to the diagnostic application developed in the LabVIEW software, National Instruments (Austin, TX, USA). After measuring the diagnostic signal, the input data were preprocessed with the use of the Matlab software. It should be emphasized that the information on actual drive system speed and load torque, visible in Figure 2a, was not used in the CNN-based detector; these signals were measured for monitoring the proper training process only.

#### 3.2. Description of the Input Data Preprocessing for the Developed CNNs

In a diagnostic system using CNN in the procedure of extracting fault symptoms, the measured input signals must be properly processed. The principle of CNN operation forces the appropriate adaptation of a network input matrix. The developed input batch should be a compromise between the size of the input matrix and the amount of information delivered to the network. If the size of the input matrix is too large, it will significantly increase network training time and, consequently, when it is too small, it will not ensure the proper division of the input data into classes.

The following steps of the data preprocessing algorithm are illustrated in Figure 3. In the first step {1}, the phase currents of the IM were measured. The research assumed the possibility of fault detection after measuring only 2000 samples of the diagnostic signal, which constituted two full periods for the minimal frequency of the motor supply voltage equal to 10 Hz.

Then, phase current vectors, with sizes  $1 \times 2000$  each, were normalized to the rated current of the tested motor {2}. In the third step, the conversion of each normalized vector (expressed in relative units) to a  $40 \times 50$  matrix was used. This operation was performed for each of the measured phase currents {3}. The last step in this procedure consisted in the conversion of three matrices containing standardized phase current samples to a three-dimensional matrix {4}. The presented method of diagnostic signal preprocessing allowed one to generate 3200 input vectors for the neural network. It should also be emphasized that the use of such a small number of samples, compared to those used in the literature [48], allowed to limit the time of a single measurement to 0.2 s.



Induction Motor

**Figure 3.** Elaboration of the input layer of the convolutional neural network (CNN) used for electrical fault detection of the induction motor—schematic diagram;  $I_{MEAS}$ —measurement value of stator currents,  $I_N$ —nominal value of stator currents.

#### 3.3. Description of the Structure and Tasks of the Developed CNN-Base Fault Detectors

The experimental verification of the developed CNNs was based on 3200 vectors of stator currents containing data on the state of the IM for 16 fault categories of the stator windings (1—for undamaged motor, 15—for the motor with a different number of the inter-turn short circuits of the stator winding (1–5 shorted turns in three stator phases). The experimental study was carried out in four stages:

- analysis of the assessment effectiveness of the degree of damage to a single stator phase of the IM—CNN-1,
- analysis of the assessment effectiveness of the degree of damage in three phases of the stator the IM—CNN-2,
- examination of the impact of CNNs' learning process parameters on the effectiveness of the IM damage assessment,
- analysis of the impact of CNNs' structure on the effectiveness of the IM damage assessment.
- The analysis of the results is presented below and in the next section of the paper.

The NN training process was performed using the Matlab environment. The basic quantities describing the developed structures and the parameters of the CNN training process are listed in Table 1.

In the research, two structures of convolutional networks, CNN-1 and CNN-2, were used to classify the degree of damage to the stator winding of the IM, based on the information about the phase currents. The differences in the structures of the discussed networks result from the task assigned to them (assessment of the number of categories). The extension of the CNN-1 structure with an additional convolutional layer gave the network the ability to recognize in which phase of the stator winding a damage occurred. A slight expansion of the structure allowed for a triple increase in the number of the IM stator fault categories. As shown in Table 1, the developed networks are characterized by a relatively small number of neurons compared to known convolutional structures [48]. This results from the fact that to solve some problems, it is required to use many layers detecting features, not necessarily having a large number of neurons.

After the training processes were both developed, CNNs were tested in on-line operation based on raw data of the stator current, different from those used in the learning procedures. Result of operation of the CNN-based detectors of inter-turn short circuits of the IM stator winding are demonstrated in the next section.

Parameters of the Developed CNN Structures			Parameters of the Training Process of CNNs			
Name of parameter	Value of parameter		Name of parameter	Value of parameter		
	CNN-1	CNN-2	France of parameter	CNN-1	CNN-2	
Number of CONV layers	3	4	Learning method	SGDM		
Numbers of filters in particular layers	30-60-90	40-60-80-100	Momentum coefficient	0.9	0.5	
Filters size	(3 × 3)	(6 × 6)	Initial learning rate	0	.001	
Padding method Stride	"same" (2,2)	"same" (4,4)	Number of training epochs	900	700	
Number of POOL layers	3	4	Learning rate dropping method	"piecewise"		
Pooling method	Maxii	mum	Drop period	20		
Pool size	(3 × 3)	(6 × 6)	L2—regularization	0.0001		
Padding Stride	"same" (2,2)	"same" (4,4)	Gradient threshold method	L2—norm		
Number of ReLU layers	3	4	Training matrix size	$40 \times 50 \times 3 \times 600$	$40 \times 50 \times 3 \times 1600$	
Number of BN layers	3	4	Testing matrix size	$40 \times 50 \times 3 \times 600$	$40 \times 50 \times 3 \times 1600$	
Epsilon	0.001	0.001	Validation frequency		10	
Number of DOUT layers	1	1	Shuffle method	"every-epoch"		
Probability	0.5	0.5	Mini batch size	80	90	
Number of FC layers	1	1	Number of categories	6	16	
Number of FC neurons	6	16	i tumber of categories	0	10	
CONV—convolutional layer BN—batch normaliza		ation layer	DOUT—dropout layer			
POOL—pooling layer ReLU—rectified linear			r unit laver	FC—fully connected laver		

Table 1. Parameters of structures and training process of the developed CNNs.

# 4. Analysis of Experimental Results of the CNN-Based Stator Winding Fault Detectors Working On-Line

The experimental verification of the developed structures was performed on the basis of the prepared test data of the following sizes: 600—for CNN-1, 1600—for CNN-2. They contained information about the measured stator currents for various degrees of stator winding faults. The test data were developed for various frequencies of the supply voltage and the load torque values of the motor. Analyzing the responses of the developed CNN structures to the test data, a very high degree of effectiveness in assessing the degree of stator winding damage can be observed.

Figure 4 presents an example of the on-line operation of the fault detector based on CNN-1. Physical damage modeling consisted of temporary shorting of five turns of phase A. This approach allowed the determination of the minimum detection time  $T_d$ , as well as the damage classification time  $T_c$ . Detection time should be understood as the time between the occurrence of a winding short-circuit and the first CNN output information, indicating damage to the machine. However, the classification time is measured from the moment the damage occurs until the system response is established, and should not be longer than the measurement time of 2000 current samples. The test results presented in Figure 4 were obtained during IM operation at various values of the supply voltage frequency and the load torque.

The analysis of the obtained results allowed one to notice that the developed neural structure allows for the correct detection of defects, even in the case of the measurement of the transient state of the machine. In addition, the machine operating conditions do not affect the precision of the diagnostic system. As can be seen in Figure 4a–d, as a result of completing the CNN input data buffer with

subsequent samples characteristic of a faulty winding, the CNN network shows a gradual change in the category of damage. The same occurrence was observed in Figure 4 during the clearance of the damage. An analysis of transient states allowed one to notice that the assessment of the level of damage is based on changes in the whole input matrix. On the other hand, damage detection is based on the outliers' samples, hence the detection time is much shorter than the classification time. Nevertheless, the main task of the diagnostic system is damage detection, while grade assessment is a secondary function.

In the next Figure 5 the accuracy of fault level assessment is demonstrated for both tested CNN-based detectors. The fault category assessment accuracy was 99.3% for the inter-turn short circuits in a single phase (Figure 5a—CNN-1) and 98.8% for the fault detection in three phases (Figure 5b—CNN-2), respectively. Errors in the assessment of the winding condition most frequently resulted from the impact of the load torque on the value of the measured current amplitudes. In Figure 5b, the individual phases of the stator winding, in which the inter-turn short-circuits occurred, were marked with different colors.



Figure 4. Cont.



**Figure 4.** CNN on-line response during motor operation: (a)  $N_{sh} = 5$ ,  $f_s = 20$  Hz,  $T_L = 0.2$   $T_N$ ,  $T_d = 0.0273$  s,  $T_c = 0.1783$  s; (b)  $N_{sh} = 5$ ,  $f_s = 20$  Hz,  $T_L = 0.6$   $T_N$ ,  $T_d = 0.0085$  s,  $T_c = 0.1985$  s; (c)  $N_{sh} = 5$ ,  $f_s = 50$  Hz,  $T_L = 0.4$   $T_N$ ,  $T_d = 0.0508$  s,  $T_c = 0.1503$  s; (d)  $N_{sh} = 5$ ,  $f_s = 50$  Hz,  $T_L = 0.8$   $T_N$ ,  $T_d = 0.022$  s,  $T_c = 0.1536$  s:  $N_{sh}$ —number of shorted turns in stator windings;  $T_N$ —nominal torque;  $T_L$ —load torque;  $f_s$ —supply voltage frequency;  $T_d$ —fault detection time;  $T_c$ —fault classification time.

The analysis of the results presented in this figure showed that CNN-2 provided a small number of incorrect responses regarding the phase in which a failure occurred (approximately 0.43%). The incorrect network responses occurred mainly at small values of the load torque. Nevertheless, both developed structures maintained a very high degree of effectiveness, although no advanced signal processing methods (FFT, DWT, HHT) were used in the diagnostic procedure. The fact that the raw diagnostic signal (measured stator current in one or three phases) went directly to the input of the CNN is, according to the authors' knowledge, a novelty in the field of the diagnostics of inter-turn short circuits of electrical machines. Additionally, the use of analytical diagnostic methods requires a few seconds of signal measurement. In the presented CNN application, the input vector was based on the current measurement in a time interval of maximum 0.2 s, however the detection and classification

times are usually smaller, as can be seen in examples of Figure 4. The differences in damage detection and classification times that can be seen in waveforms shown in Figure 4 are due to the fact that with the variable operating conditions of the drive (frequency or load change), the NN needs more or less samples to determine the damage class correctly. Therefore, the proposed CNN structures allow the multiple shortening of the detection process, while maintaining a high efficiency level, which due to the avalanche spreading nature of the analyzed damages, such as inter-turn short-circuits of the stator winding, is an undoubted advantage.



**Figure 5.** Confusion matrix for testing data: (a) CNN-1; (b) CNN-2: red envelope—phase A; blue envelope—phase B; green envelope—phase C; *N*<sub>sh</sub>—number of shorted turns in stator windings.

## 5. Impact of the CNN Structure and Training Method Parameters to the Fault Diagnosis

# 5.1. Impact of the Convolutional Network Training Parameters on the Effectiveness of IM Stator Damage Assessment

The appropriate selection of the network structure for the analyzed task can ensure high CNN precision only with the properly adjusted parameters of the training process. The research focused on the simple gradient algorithm, which is the most popular method in the deep network training [62]. Detailed studies on the impact of training process parameters were carried out for the developed CNN-1 and CNN-2 structures and the results are shown in Figure 6.

According to the SGDM algorithm, the right choice of the data packet size, momentum factor and the frequency of parameter update (number of epochs, number of iterations in the epoch) is crucial.

As observed in Figure 6, the impact of the training parameters on the effectiveness of individual structures is similar. The differences result from a larger number of layers in CNN-2, as well as from a much broader learning data vector. The largest differences in the accuracy of the neural detector are noticeable when the momentum factor changes. The CNN-1 structure, with a much smaller number of classes required for recognition, is characterized by higher efficiency for much higher values of the momentum coefficient (Figure 6c). The high value of this parameter during the training of the CNN-1 network resulted in a rapid decrease in the value of the loss function, and at the same time an increase in the number of learning epochs necessary to ensure maximum precision (Figure 6a). The analysis of Figure 6b shows that the size of the data batch significantly affects the training process of both developed structures. The research abandoned the standard approach to choosing the batch size as a power of two. This approach results from the fact that memory operations are optimized for processing 2n arrays, while the structures presented here did not require such accurate optimization of the training process.

The selection of structures, as well as the parameters of the training process in accordance with this principle, can be justified by much more complex CNN structures. The right number of network parameter updates becomes much more important for the training process. This property is directly related to the number of learning epochs and the size of the data batch. The number of network parameter updates with a fixed number of epochs decreases as the batch packet size increases. Therefore, as the batch size increases, the number of learning epochs should be increased accordingly. The number of parameter updates equal to the quotient of the training vector size and the data packet size should always be constant. A very important technique used to improve the training process is the shuffling of the validation and training data. As can be seen in Figure 6d, the structures in which shuffling was used are characterized by a much higher level of efficiency. The shuffling method before each training epoch allows you to achieve the highest level of effectiveness. When the size of the training data batch does not contain an even number of samples from each category, some samples are discarded. The use of data shuffling before each epoch prevents the situation in which certain data is regularly discarded in each epoch.



**Figure 6.** The impact of training parameters on the effectiveness of CNN structures: (**a**) the number of training periods; (**b**) data packet size; (**c**) momentum factor; (**d**) shuffling method.

### 5.2. Impact of the Convolutional Network Structure on the Effectiveness of IM Stator Damage Assessment

The task of assessing the technical condition of electrical machines operating in converter-fed drives imposes the additional requirements for neural systems such as a high level of detection efficiency, high accuracy in class evaluation, insensitivity to measurement disturbances, and the ability to generalize. Meeting these requirements is possible only with the appropriate selection of the network structure, learning methods, as well as the appropriate choice of the diagnostic information carrier. As presented in the previous section, changes in the network structure affect changes in training process parameters.

In order to optimize the process of selecting network parameters as much as possible, one should simultaneously examine the impact of changes in the structure, as well as in the learning process. In Table 2, a summary of various CNN-1 structure configurations is presented to identify the relationship between the number of layers used and the network efficiency. This table has been divided into five sections, due to different changes made in the network structure. The bold fonts used in this table show what is changed in the CNN structure in the conducted experiments; e.g., **Activation function** shows, that this finction has been changed, while the rest of CNN's parameters were the same in the presented experiment. Also the best obtained accuracy has been marked with bold font.

Number of Layers and Other Parameters						
Convolutional Layer	Activation Function	Pooling Layer	Batch Normalization Layer	Dropout Layer	Fully Connected Layer	Accuracy
4	4	4	4	1	1	97.0%
30-60-90-120	ReLU	Pooling Max	Epsilon = 0.0001	Probability = 0.5	Output Size = 6	
3	3	3	3	1	1	99.3%
30-60-90	ReLU	Pooling Max	Epsilon = 0.0001	Probability = 0.5	Output Size = 6	
2	2	2	2	1	1	98.0%
30-60-90	ReLU	Pooling Max	Epsilon = 0.0001	Probability = 0.5	Output Size = 6	
1	1	1	1	1	1	96.5%
30-60-90	ReLU	Pooling Max	Epsilon = 0.0001	Probability = 0.5	Output Size = 6	
Convolutional Layer	Activation Function	Pooling Layer	Batch Normalization Layer	Dropout Layer	Fully Connected Layer	Accuracy
3	3	3	3	1	1	99.3%
30-60-90	ReLU	Pooling Max	Epsilon = 0.0001	Probability = 0.5	Output Size = 6	
3 30-60-90	3 Hyperbolic Tangent	3 Pooling Max	3 Epsilon = 0.0001	1 Probability = 0.5	1 Output Size = 6	96.8%
3	3	3	3	1	1	97.5%
30-60-90	Leaky ReLU	Pooling Max	Epsilon = 0.0001	Probability = 0.5	Output Size = 6	
3	3	3	3	1	1	98.2%
30-60-90	Clipped ReLU	Pooling Max	Epsilon = 0.0001	Probability = 0.5	Output Size = 6	
3 30-60-90	3 Exp. Linear Unit	3 Pooling Max	3 Epsilon = 0.0001	1 Probability = 0.5	1 Output Size = 6	98.2%

Table 2. The effect of the parameters of the neural structure on the effectiveness of the CNN-based damage detected	or.
---	-----

Number of Layers and Other Parameters						
Convolutional Layer	Activation Function	Pooling Layer	Batch Normalization Layer	Dropout Layer	Fully Connected Layer	Accuracy
3	3	3	3	1	1	99.3%
30-60-90	ReLU	Pooling Max	Epsilon = 0.0001	Probability = 0.5	Output Size = 6	
3	3	3	3	1	1	98.7%
30-60-90	ReLU	Pooling Average	Epsilon = 0.0001	Probability = 0.5	Output Size = 6	
3	3	2	3	1	1	98.5%
30-60-90	ReLU	Pooling Max	Epsilon = 0.0001	Probability = 0.5	Output Size = 6	
3	3	2	3	1	1	97.7%
30-60-90	ReLU	Pooling Average	Epsilon = 0.0001	Probability = 0.5	Output Size = 6	
3	3	1	3	1	1	97.6%
30-60-90	ReLU	Pooling Max	Epsilon = 0.0001	Probability = 0.5	Output Size = 6	
3	3	1	3	1	1	96.5%
30-60-90	ReLU	Pooling Average	Epsilon = 0.0001	Probability = 0.5	Output Size = 6	
Convolutional Layer	Activation Function	Pooling Layer	Batch Normalization Layer	Dropout Layer	Fully Connected Layer	Accuracy
3	3	3	3	1	1	99.3%
30-60-90	ReLU	Pooling Max	Epsilon = 0.0001	Probability = 0.5	Output Size = 6	
3	3	3	2	1	1	97.7%
30-60-90	ReLU	Pooling Max	Epsilon = 0.0001	Probability = 0.5	Output Size = 6	
3	3	3	1	1	1	97.8%
30-60-90	ReLU	Pooling Max	Epsilon = 0.0001	Probability = 0.5	Output Size = 6	

Number of Layers and Other Parameters						
Convolutional Layer	Activation Function	Pooling Layer	Batch Normalization Layer	Dropout Layer and Fully	Connected Layer	Accuracy
3	3	3	3	1	1	99.3%
30-60-90	ReLU	Pooling Max	Epsilon = 0.0001	Probability = 0.5	Output Size = 6	
3	3	3	3	1	2	99.3%
30-60-90	ReLU	Pooling Max	Epsilon = 0.0001	Probability = 0.5	Structure: 36-6	
3	3	3	3	1	3	99.3%
30-60-90	ReLU	Pooling Max	Epsilon = 0.0001	Probability = 0.5	Structure: 72-36-6	
3	3	3	3	0	1	94.8%
30-60-90	ReLU	Pooling Max	Epsilon = 0.0001	Probability = 0.5	Output Size = 6	
3	3	3	3	0	2	95.7%
30-60-90	ReLU	Pooling Max	Epsilon = 0.0001	Probability = 0.5	Structure: 36-6	
3	3	3	3	0	3	94.7%
30-60-90	ReLU	Pooling Max	Epsilon = 0.0001	Probability = 0.5	Structure: 72-36-6	
Used description of the convolutional layer		3	Number of convolutional layers		yers	
		30-60-90	Number of filters in 1st, 2nd, 3th layer		h layer	

Table 2. Cont.

In the first step, the effect of the number of convolutional layers was analyzed. It was noted that starting from the basic structure of CNN-1 with a fixed number of categories necessary to recognize, placing an additional convolutional layer does not increase the network efficiency. However, an additional layer significantly increased the number of network parameters, which made the training process longer.

Subsequent changes in the CNN structure concerned the activation function. The undoubted advantage of the ReLU (rectified linear unit layer) function used together with SGDM is a significant increase in the dynamics of the training process, compared to the function of the hyperbolic tangent. The analysis of the results presented in the second section of Table 2 shows the superiority of the ReLU function over the other activation functions.

The use of pooling layers in the CNN structure makes the network more resistant to noise. This fact results from the principle of operation of this layer which provides information about the maximum or average values of a given window. The combination of convolutional and pooling layers results in a significant increase in the network precision, which can be observed by analyzing the results from the third section of Table 2. The advantage of the maximum search method over the calculation of the average value of the window is also noticeable. A clear improvement in CNN performance is visible when using packet normalization at the output of each of the convolutional layers (see the fourth section of Table 2). Thanks to this approach, the mapping of features is independent of their spatial location, which directly affects the network efficiency, especially with test data being significantly different from training packages.

The use of the dropout layer results in an increase in training time and also allows one to obtain more generalized models with higher precision, which can be seen in the fifth section of Table 2.

The research referred to one level of probability equal to 0.5, which results from the lack of influence of this parameter on the network effectiveness. The dropout layer was applied only before the application of fully connected layers, so as not to miss a part of the input data set. The analysis of the fifth section of Table 2 also shows that the use of multiple fully connected layers to better determine the impact of individual features does not have the intended effect. Therefore, it is enough to use such a layer with the number of neurons equal to the number of recognized categories, in combination with classifying layers.

### 6. Conclusions

The main goal of the research presented in this paper has been achieved. It has been proven by experimental tests that the application of the CNN, with a relatively simple structure compared to those used in the literature, for the detection and classification of the IM incipient stator winding faults, offers interesting results. The proper detection of the incipient inter-turn short circuits of the IM stator winding was achieved after measuring maximum 2000 samples of the diagnostic signal-stator current transient. It should also be emphasized that the use of such a small number of samples, compared to those used in the literature, allowed one to limit the maximum time of stator winding fault detection to a maximum of 0.2 s, which is directly connected with the size of the input matrix. Nevertheless, the on-line tests show that the high accuracy of fault detection is obtained, based on less than 200 samples equivalent to 0.02 s of measurement. Moreover, the detection system responses are not dependent on the motor operating conditions. The impact of changes in the structure of the analyzed CNNs, as well as parameters of the learning process, have been analyzed in the presented research and the detailed discussion was presented in the fourth section of this article. The two developed structures of CNNs, one for the assessment of the degree of damage to a single stator phase and the other one for assessing the degree of damage to all three phases, maintain a very high degree of effectiveness, based on the direct processing of the raw measured stator current data, without the use of advanced signal processing methods (FFT, DWT, HHT) in the diagnostic procedure. This is, according to the best of the authors' knowledge, a novelty in the fault diagnostic and classification of the stator winding faults using NNs, not only shallow but also deep learning structures.

The main advantages of the presented IM stator winding fault detectors based on CNN networks and stator current measurement are:

- the use of raw measurement data as network input signals,
- no need to pre-process measurement data with analytical methods,
- high accuracy in detecting the localization of the failure (motor phase) and the fault level (number of shorted turns),
- the ability to detect even individual shorted turns (incipient faults),
- a merger of symptom extraction and damage detection in one machine learning process.

It has been also shown, that after proper training, the developed CNNs can be used for on-line fault detection and their level assessment using program cooperation between Matlab and LabVIEW software. Future work of the authors will be focused on on-line implementation of CNN-based stator winding incipient fault detectors using microcontroller systems and FPGA.

**Author Contributions:** All the authors contributed equally to the concept of the paper, and proposed the methodology, M.S. and M.W. carried out the measurements and analyzed the experimental data, M.S. designed and implemented the CNNs detectors, M.S. and T.O.-K. proposed the paper organization, prepared the original draft, C.T.K. validated the obtained results, T.O.-K. Supervised the project, revised and edited the final paper. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was partly supported by the National Science Centre Poland under grant number 2017/27/B/ST7/00816 and by statutory funds of the Department of Electrical Machines, Drives and Measurements of the Wroclaw University of Science and Technology (2020).

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

### Appendix A

Name of the Parameter	Symbol	Units	
Power	$P_N$	3000	[W]
Torque	$T_N$	19.83	[Nm]
Speed	$N_N$	1445	[r/min]
Stator phase voltage	$U_{sN}$	230	[V]
Stator current	$I_{sN}$	6.8	[A]
Frequency	$f_{sN}$	50	[Hz]
Pole pairs number	$p_p$	2	[-]

Table A1. Rated parameters of the tested induction motor.

### Appendix **B**



**Figure A1.** Illustration of the physical modelling of stator fault: (**a**) connection of the stator winding for inter-turn short-circuit modelling; (**b**) tested induction motor—real view.

### References

- 1. Stone, G.; Boutler, E.; Cubert, I.; Dhirani, H. *Electrical Insulation for Rotating Machines- Design, Evaluation, Aging, Testing, and Repair,* 2nd ed.; John Wiley & Sons: Hoboken, NJ, USA, 2014.
- 2. Benbouzid, M.E.H.; Vieira, M.; Theys, C. Induction motors faults detection and localization using stator current advanced signal processing techniques. *IEEE Trans. Power Electron.* **1999**, *14*, 14–22. [CrossRef]
- 3. Schoen, R.R.; Habetler, T.G.; Kamran, F.; Bartfield, R.G. Motor bearing damage detection using stator current monitoring. *IEEE Trans. Ind. Appl.* **1995**, *31*, 1274–1279. [CrossRef]
- 4. Fang, J.; Sun, Y.; Wang, Y.; Wei, B.; Hang, J. Improved ZSVC-based fault detection technique for incipient stage inter-turn fault in PMSM. *IET Electric Power Appl.* **2019**, *13*, 2015–2026. [CrossRef]
- Urresty, J.; Riba, J.; Romeral, L. A Back-emf Based Method to Detect Magnet Failures in PMSMs. *IEEE Trans.* Magn. 2013, 49, 591–598. [CrossRef]
- Garcia-Perez, A.; Romero-Troncoso, R.J.; Cabal-Yepez, E.; Osornio-Rios, R.A. The Application of High-Resolution Spectral Analysis for Identifying Multiple Combined Faults in Induction Motors. *IEEE Trans. Ind. Electron.* 2011, 58, 2002–2010. [CrossRef]
- 7. Harmouche, J.; Delpha, C.; Diallo, D. Improved fault diagnosis of ball bearings based on the global spectrum of vibration signals. *IEEE Trans. Energy Convers.* **2015**, *30*, 376–383. [CrossRef]
- 8. Ewert, P. Use of axial flux in the detection of electrical faults in induction motors. In Proceedings of the 2017 International Symposium on Electrical Machines (SME), Nałęczów, Poland, 18–21 June 2017. [CrossRef]
- Skowron, M.; Wolkiewicz, M.; Orlowska-Kowalska, T.; Kowalski, C.T. Effectiveness of Selected Neural Network Structures Based on Axial Flux Analysis in Stator and Rotor Winding Incipient Fault Detection of Inverter-fed Induction Motors. *Energies* 2019, 12, 2392. [CrossRef]
- Picazo-Ródenas, M.J.; Antonino-Daviu, J.; Climente-Alarcon, V.; Royo-Pastor, R.; Mota-Villar, A. Combination of Noninvasive Approaches for General Assessment of Induction Motors. *IEEE Trans. Ind. Appl.* 2015, 51, 2172–2180. [CrossRef]
- Kumar, P.S.; Xie, L.; Halick, M.S.M.; Vaiyapuri, V. Online stator end winding thermography using infrared sensor array. In Proceedings of the 2018 IEEE Applied Power Electronics Conference and Exposition (APEC), San Antonio, TX, USA, 4–8 March 2018. [CrossRef]
- 12. Nejjari, H.; Benbouzid, M.E.H. Monitoring and diagnosis of induction motors electrical faults using a current Park's vector pattern learning approach. *IEEE Trans. Ind. Appl.* **2000**, *36*, 730–735. [CrossRef]
- Puche-Panadero, R.; Pineda-Sanchez, M.; Riera-Guasp, M.; Roger-Folch, J.; Hurtado-Perez, E.; Perez-Cruz, J. Improved Resolution of the MCSA Method Via Hilbert Transform, Enabling the Diagnosis of Rotor Asymmetries at Very Low Slip. *IEEE Trans. Energy Convers.* 2009, 24, 52–59. [CrossRef]
- Arabaci, H.; Bilgin, O. The Detection of Rotor Faults By Using Short Time Fourier Transform. In Proceedings of the 2007 IEEE 15th Signal Processing and Communications Applications, Eskisehir, Turkey, 11–13 June 2007. [CrossRef]
- 15. Jung, J.; Lee, J.; Kwon, B. Online Diagnosis of Induction Motors Using MCSA. *IEEE Trans. Ind. Electron.* **2006**, 53, 1842–1852. [CrossRef]
- Haddad, R.Z.; Strangas, E.G. On the Accuracy of Fault Detection and Separation in Permanent Magnet Synchronous Machines Using MCSA/MVSA and LDA. *IEEE Trans. Energy Convers.* 2016, 31, 924–934. [CrossRef]
- 17. Cruz, S.M.A.; Cardoso, A.J.M. Stator winding fault diagnosis in three-phase synchronous and asynchronous motors, by the extended Park's vector approach. *IEEE Trans. Ind. Appl.* **2016**, *37*, 1227–1233. [CrossRef]
- Rosero, J.; Cusido, J.; Espinosa, A.G.; Ortega, J.A.; Romeral, L. Broken Bearings Fault Detection for a Permanent Magnet Synchronous Motor under non-constant working conditions by means of a Joint Time Frequency Analysis. In Proceedings of the 2007 IEEE International Symposium on Industrial Electronics, Vigo, Spain, 4–7 June 2007. [CrossRef]
- Espinosa, A.G.; Rosero, J.A.; Cusidó, J.; Romeral, L.; Ortega, J.A. Fault Detection by Means of Hilbert–Huang Transform of the Stator Current in a PMSM With Demagnetization. *IEEE Trans. Energy Convers.* 2010, 25, 312–318. [CrossRef]

- Rosero, J.; Garcia, A.; Cusido, J.; Romeral, L.; Ortega, J. Fault detection by means of Hilbert Huang Transform of the stator current in a PMSM with demagnetization. In Proceedings of the 2007 IEEE International Symposium on Intelligent Signal Processing, Alcala De Henares, Spain, 3–5 October 2007. [CrossRef]
- 21. Ruiz, J.R.; Rosero, J.A.; Espinosa, A.G.; Romeral, L. Detection of Demagnetization Faults in Permanent-Magnet Synchronous Motors Under Nonstationary Conditions. *IEEE Trans. Magn.* 2009, 45, 2961–2969. [CrossRef]
- 22. Park, C.H.; Lee, J.; Ahn, G.; Youn, M.; Youn, B.D. Fault Detection of PMSM under Non-Stationary Conditions Based on Wavelet Transformation Combined with Distance Approach. In Proceedings of the 2019 IEEE 12th International Symposium on Diagnostics for Electrical Machines, Power Electronics and Drives (SDEMPED), Toulouse, France, 27–30 August 2019. [CrossRef]
- 23. Riera-Guasp, M.; Antonino-Daviu, J.A.; Pineda-Sanchez, M.; Puche-Panadero, R.; Perez-Cruz, J. A General Approach for the Transient Detection of Slip-Dependent Fault Components Based on the Discrete Wavelet Transform. *IEEE Trans. Ind. Electron.* **2008**, *55*, 4167–4180. [CrossRef]
- Abdellatif, S.; Aissa, C.; Hamou, A.A.; Chawki, S.; Oussama, B.S. A Deep Learning Based on Sparse Auto-Encoder with MCSA for Broken Rotor Bar Fault Detection and Diagnosis. In Proceedings of the 2018 International Conference on Electrical Sciences and Technologies in Maghreb (CISTEM), Alger, Algeria, 28–31 October 2018. [CrossRef]
- 25. Ince, T.; Kiranyaz, S.; Eren, L.; Askar, M.; Gabbouj, M. Real-Time Motor Fault Detection by 1-D Convolutional Neural Networks. *IEEE Trans. Ind. Electron.* **2016**, *63*, 7067–7075. [CrossRef]
- 26. Junbo, T.; Weining, L.; Juneng, A.; Xueqian, W. Fault diagnosis method study in roller bearing based on wavelet transform and stacked auto-encoder. In Proceedings of the 27th Chinese Control and Decision Conference (2015 CCDC), Qingdao, China, 23–25 May2015. [CrossRef]
- 27. Lee, Y.O.; Jo, J.; Hwang, J. Application of deep neural network and generative adversarial network to industrial maintenance: A case study of induction motor fault detection. In Proceedings of the 2017 IEEE International Conference on Big Data (Big Data), Boston, MA, USA, 11–14 December 2017. [CrossRef]
- 28. Hamdani, S.; Touhami, O.; Ibtiouen, R.; Fadel, M. Neural network technique for induction motor rotor faults classification-dynamic eccentricity and broken bar faults. In Proceedings of the 8th IEEE Symposium on Diagnostics for Electrical Machines, Power Electronics Drives, Bologna, Italy, 5–8 September 2011. [CrossRef]
- 29. Ghate, V.N.; Dudul, S.V. Optimal MLP neural network classifier for fault detection of three phase induction motor. *Expert Syst. Appl.* **2010**, *37*, 3468–3481. [CrossRef]
- 30. Taïbi, Z.M.; Hasni, M.; Hamdani, S.; Rahmani, O.; Touhami, O.; Ibtiouen, R. Optimization of the feedforward neural network for rotor cage fault diagnosis in three-phase induction motors. In Proceedings of the 2011 IEEE International Electric Machines Drives Conference (IEMDC), Niagara Falls, ON, Canada, 15–18 May 2011. [CrossRef]
- 31. Skowron, M.; Wolkiewicz, M.; Orlowska-Kowalska, T.; Kowalski, C.T. Application of Self-Organizing Neural Networks to Electrical Fault Classification in Induction Motors. *Appl. Sci.* **2019**, *9*, 616. [CrossRef]
- 32. Kowalski, C.T.; Orlowska-Kowalska, T. Neural networks application for induction motor faults diagnosis. *Math. Comput. Simul.* **2003**, 63, 435–448. [CrossRef]
- Khalfaoui, N.; Salhi, M.S.; Amiri, H. The SOM tool in mechanical fault detection over an electric asynchronous drive. In Proceedings of the 2016 4th International Conference on Control Engineering Information Technology (CEIT), Hammamet, Tunisia, 16–18 December 2016. [CrossRef]
- Sid, O.; Menaa, M.; Hamdani, S.; Touhami, O.; Ibtiouen, R. Self-organizing map approach for classification of electricals rotor faults in induction motors. In Proceedings of the 2011 2nd International Conference on Electric Power and Energy Conversion Systems (EPECS), Niagara Falls, ON, Canada, 15–18 May 2011. [CrossRef]
- Afrasiabi, S.; Afrasiabi, M.; Parang, B.; Mohammadi, M. Real-Time Bearing Fault Diagnosis of Induction Motors with Accelerated Deep Learning Approach. In Proceedings of the 2019 10th International Power Electronics, Drive Systems and Technologies Conference (PEDSTC), Shiraz, Iran, 12–14 February 2019. [CrossRef]
- 36. Shao, S.; Yan, R.; Lu, Y.; Wang, P.; Gao, R. DCNN-based Multi-signal Induction Motor Fault Diagnosis. *IEEE Trans. Instrum. Meas.* **2019**. [CrossRef]

- 37. Shao, S.; McAleer, S.; Yan, R.; Baldi, P. Highly Accurate Machine Fault Diagnosis Using Deep Transfer Learning. *IEEE Trans. Ind. Inform.* **2019**, *15*, 2446–2455. [CrossRef]
- 38. Chen, Z.; Gryllias, K.; Li, W. Mechanical fault diagnosis using Convolutional Neural Networks and Extreme Learning Machine. *Mech. Syst. Signal Process.* **2019**, *133*, 106272. [CrossRef]
- 39. Lee, J.-H.; Pack, J.-H.; Lee, I.-S. Fault Diagnosis of Induction Motor Using Convolutional Neural Network. *Appl. Sci.* **2019**, *9*, 2950. [CrossRef]
- 40. Li, X.; Zhang, W.; Ding, Q.; Sun, J.-Q. Multi-Layer domain adaptation method for rolling bearing fault diagnosis. *Signal Process.* **2019**, *157*, 180–197. [CrossRef]
- 41. Sun, W.; Zhao, R.; Yan, R.; Shao, S.; Chen, X. Convolutional Discriminative Feature Learning for Induction Motor Fault Diagnosis. *IEEE Trans. Ind. Inform.* **2017**, *13*, 1350–1359. [CrossRef]
- 42. Kao, I.; Wang, W.; Lai, Y.; Perng, J. Analysis of Permanent Magnet Synchronous Motor Fault Diagnosis Based on Learning. *IEEE Trans. Instrum. Meas.* **2019**, *68*, 310–324. [CrossRef]
- Khan, T.; Alekhya, P.; Seshadrinath, J. Incipient Inter-turn Fault Diagnosis in Induction motors using CNN and LSTM based Methods. In Proceedings of the 2018 IEEE Industry Applications Society Annual Meeting (IAS), Portland, OR, USA, 23–27 September 2018. [CrossRef]
- 44. Hoang, D.T.; Kang, H.J. A Motor Current Signal Based Bearing Fault Diagnosis Using Deep Learning And Information Fusion. *IEEE Trans. Instrum. Meas.* **2019**. [CrossRef]
- 45. Pandarakone, S.E.; Masuko, M.; Mizuno, Y.; Nakamura, H. Deep Neural Network Based Bearing Fault Diagnosis of Induction Motor Using Fast Fourier Transform Analysis. In Proceedings of the 2018 IEEE Energy Conversion Congress and Exposition (ECCE), Portland, OR, USA, 23–27 September 2018. [CrossRef]
- 46. Xu, G.; Liu, M.; Jiang, Z.; Shen, W.; Huang, C. Online Fault Diagnosis Method Based on Transfer Convolutional Neural Networks. *IEEE Trans. Instrum. Meas.* **2020**, *69*, 509–520. [CrossRef]
- 47. Xu, G.; Liu, M.; Jiang, Z.; Söffker, D.; Shen, W. Bearing Fault Diagnosis Method Based on Deep Convolutional Neural Network and Random Forest Ensemble Learning. *Sensors* **2019**, *19*, 1088. [CrossRef]
- 48. Wen, L.; Li, X.; Gao, L.; Zhang, Y. A New Convolutional Neural Network-Based Data-Driven Fault Diagnosis Method. *IEEE Trans. Ind. Electron.* **2018**, *65*, 5990–5998. [CrossRef]
- Chattopadhyay, P.; Saha, N.; Delpha, C.; Sil, J. Deep Learning in Fault Diagnosis of Induction Motor Drives. In Proceedings of the 2018 Prognostics and System Health Management Conference (PHM-Chongqing), Chongqing, China, 26–28 October 2018. [CrossRef]
- 50. Li, C.; Zhang, W.; Peng, G.; Liu, S. Bearing Fault Diagnosis Using Fully-Connected Winner-Take-All Autoencoder. *IEEE Access* 2018, *6*, 6103–6115. [CrossRef]
- 51. Principi, E.; Rossetti, D.; Squartini, S.; Piazza, F. Unsupervised electric motor fault detection by using deep autoencoders. *IEEE/CAA J. Autom. Sin.* 2019, *6*, 441–451. [CrossRef]
- 52. Ding, X.; He, Q. Energy-Fluctuated Multiscale Feature Learning With Deep ConvNet for Intelligent Spindle Bearing Fault Diagnosis. *IEEE Trans. Instrum. Meas.* **2017**, *66*, 1926–1935. [CrossRef]
- 53. Gardel, P.; Morinigo-Sotelo, D.; Duque-Perez, O.; Perez-Alonso, M.; Garcia-Escudero, L.A. Neural network broken bar detection using time domain and current spectrum data. In Proceedings of the 2012 XXth International Conference on Electrical Machines, Marseille, France, 2–5 September 2012. [CrossRef]
- 54. Krizhevsky, A.; Sutskever, I.; Hinton, G. ImageNet classification with deep convolutional neural networks. *Adv. Neural Inf. Process. Syst.* **2012**, *25*, 1106–1114. [CrossRef]
- Kalchbrenner, N.; Grefenstette, E.; Blunsom, P. A convolutional neural network for modelling sentences. In Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics, Baltimore, MD, USA, 23–25 June 2014.
- 56. Ioffe, S.; Szegedy, C. Batch normalization: Accelerating deep network training by reducing internal covariate shift. *arXiv* **2015**, arXiv:1502.03167.
- 57. Nair, V.; Hinton, G.E. Rectified linear units improve restricted Boltzmann machines. In Proceedings of the 27th International Conference on Machine Learning, Haifa, Israel, 21–24 June 2010.
- 58. Graham, B. Fractional max-pooling. arXiv 2014, arXiv:1412.6071.
- 59. Wang, S.; Manning, C. Fast dropout training. In Proceedings of the International Conference on Machine Learning, Atlanta, GA, USA, 16–21 June 2013. [CrossRef]
- 60. Srivastava, N.; Hinton, G.; Krizhevsky, A.; Sutskever, I.; Salakhutdinov, R. Dropout: A Simple Way to Prevent Neural Networks from Overfitting. *J. Mach. Learn. Res.* **2014**, *15*, 1929–1958. [CrossRef]

- 61. Yuan, B. Efficient hardware architecture of softmax layer in deep neural network. In Proceedings of the 29th IEEE International System-on-Chip Conference (SOCC), Seattle, WA, USA, 6–9 September 2016; pp. 323–326. [CrossRef]
- Dogo, E.M.; Afolabi, O.J.; Nwulu, N.I.; Twala, B.; Aigbavboa, C.O. A Comparative Analysis of Gradient Descent-Based Optimization Algorithms on Convolutional Neural Networks. In Proceedings of the 2018 Int. Conference on Computational Techniques, Electronics and Mechanical Systems (CTEMS), Belgaum, India, 21–22 December 2018; pp. 92–99. [CrossRef]



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).