Application of deep learning neural networks for the diagnosis of electrical damage to the induction motor using the axial flux

M. SKOWRON*

Department of Electrical Machines, Drives and Measurements, Wroclaw University of Science and Technology, 50-370 Wroclaw, Poland

Abstract. In industrial drive systems, one of the widest group of machines are induction motors. During normal operation, these machines are exposed to various types of damages, resulting in high economic losses. Electrical circuits damages are more than half of all damages appearing in induction motors. In connection with the above, the task of early detection of machine defects becomes a priority in modern drive systems. The article presents the possibility of using deep neural networks to detect stator and rotor damages. The opportunity of detecting shorted turns and the broken rotor bars with the use of an axial flux signal is presented.

Key words: induction motor; axial flux; deep learning; convolutional neural networks.

1. Introduction

During the operation of drive systems based on the electrical motors, various types of damage may occur, which results from the incorrect operating conditions or design defects. Currently, the modern production lines are designed to minimize the risk of unplanned downtime due to the failures that can cause losses, often exceeding the cost of purchasing new machines. Therefore, there is great need for diagnostic systems that detect the damage at the earliest possible stage. The currently used electrical machines diagnostic applications operate based on the analysis of the signals available on the object or calculated based on mathematical models [1–3].

The use of diagnostics based on the signal analysis relates to the necessity to meet the requirements, among others: maintaining the motor stationarity during the measurement, adjusting signal parameters to the method of its processing. On the other hand, the analytical methods require detailed knowledge of the structure of the tested object and its parameters. Another limitation of these methods is the fact that the role of an expert in such systems is played by a man, which results in extended analysis time, as well as lack of automation of the detection process. To limit the role of the expert in the fully automated diagnostic systems, artificial intelligence methods are used, especially artificial neural networks [3–7].

Among the neuronal fault detectors used in the field of the electric motor diagnostics for the special attention deserve:

- multilayer perceptron (MLP) [3–5],
- self-organizing Kohonen maps (SOM) [8, 9],
- radial basis function neural network (RBF) [7, 10, 11].

In the vast majority of cases, motor fault detection systems use the multilayer perceptron trained based on the data obtained from the analysis of the diagnostic signals by selected analytical methods, e.g. Fast Fourier Transform, Hilbert Transform, Wavelet Transform.

In recent years, there has been a dynamic development of artificial intelligence methods, mainly related to the concept of deep learning [12–16]. Deep learning neural networks, as opposed to the classical neural structures, are characterized by a very extensive topology. Thanks to the use of an increased number of neurons, hidden layers, as well as the advanced training methods, these networks acquire features unknown in neural networks used so far.

Deep neural networks have been used for a long time mainly in IT systems (image processing, speech recognition). Over the last two years, there has been a noticeable increase in DNN’s interest in the electrical motors fault detection systems. For the most part, they are related to the analysis of mechanical damage of the induction motor [12–15]. Most DNN-based systems use the measurement of mechanical vibrations [13, 15, 16], less frequently stator currents [12] or voltages [14]. Among the structures of deep learning networks, the most popular are the convolutional neural networks (CNN) [13, 14] and autoencoders [16]. The CNN used in diagnostic processes are characterized by a much higher level of accuracy compared with MLP, RBF networks, as presented, among others, in [13, 14]. The main difficulties in the use of DNN in diagnostic processes are to properly match the signal measured to the structure and properties of the network. The input vector of deep learning neural networks can be the result of signal analysis [12, 14], as well as a directly given diagnostic signal [13]. The choice of diagnostic signals has a significant impact on the network structure as well as on the parameters of the training process.

In the article, the possibility of detecting IM electrical damage with the use of the axial flux signal and the convolutional neural network is discussed. The paper is divided into 3 main parts. In the first part, the application of the axial flux signal in the process of induction motor damage detection is presented. The second part is devoted to the issue of the convolutional neu-
nal network (the idea of network layers as well as the developed structure). The last part includes the analysis of experimental verification of the developed convolutional neural structures.

2. Application of the axial flux signal in the detection of IM electrical fault

Electrical machines are characterized by the existence of certain asymmetries in the electrical or magnetic circuits. These asymmetries are mainly related to the limited accuracy of the technological processing of machine components, as well as modes of the motor operation. The effect of these imperfections is the existence of leakage fluxes, the value of which will depend on the level of motor asymmetry [17]. As the axial flux finds its source in the currents flowing through the motor windings, damage to the electrical circuits will also be reflected in this signal. The interaction between the damage and the flux value allows these physical changes to be used to assess the degree of defect in electric motors (Fig. 2).

Fig. 1. External stray (axial) flux in a squirrel-cage induction motor (a) and method of placement of axial flux measuring coils (b)

The analysis of machine condition assessment based on the axial flux signal in most applications is carried out using Fast Fourier Transform. Despite the high efficiency of damage detection, FFT requires stationarity as well as a relatively long measurement time. Measurement time plays a key role in the case of stator winding faults, which are characterized by extremely fast progression of the damage. In the Fig. 3, the influence of faults in the IM electrical circuit on the spectrum of the voltage induced in the measuring coil by the axial flux is presented. The spectra shown in Fig. 3 allow observing how the shorted turns (Fig. 3b) and broken rotor bar (Fig. 3a) affect the axial flux signal. The analysis of the axial flux is based on the observation of the amplitudes of spectrum components characteristic for the particular defects. The assessment of the damage degree will depend on the trend of changes in the amplitude values of individual harmonic components which were described, among others, in [17]:

- shorted turns in stator windings,

\[ f_{sh} = f_s \left( k \frac{(1-s)}{p_p} \pm m \right), \]  

- broken rotor bars,

\[ f_{bb1} = f_s (1 \pm 2ks), \]  
\[ f_{bb2} = m s f_s, \]  
\[ f_{bb3} = k f_s \pm m sf_s, \]

where: \( p_p \) – number of pole pairs, \( f_r \) – rotational frequency, \( m = 1, 3, 5, \ldots, 2p_p - 1, k = 1, 2, 3, \ldots \),

The advantage of using the axial flux signal in diagnostic processes is the non-invasiveness of the measurement, low cost of measuring systems, and high sensitivity to magnetic asymmetries occurring during damage. However, the need for complex processing of the measurement signal as well as the extraction and evaluation of the amplitudes of the characteristic damage frequencies are related to the extended measurement.
time. Therefore, to significantly reduce the time of detection and damage assessment, the method of direct signal processing is proposed.

3. Convolutional neural networks

The convolutional neural networks are the main representatives of deep neural networks. They are characterized by an expanded architecture as well as advanced training methods. The primary function of CNN is to extract the characteristic features from the input information with the use of operation of the mathematical convolution. In the presented application of CNN in the diagnostic process, the first layer can be understood as a filter of basic features. The subsequent execution of the convolution operation allows detecting the characteristics features of a higher order. Therefore, the structure of the network will depend on the type of information provided as well as the task of CNN. In the aim of detection of the complex features in the input matrix, the structures containing a lot of the convolution layers are used.

3.1. Diagnostic signal processing. In the diagnostic systems based on CNN during the process of extraction of the damage symptoms, it is necessary to convert the measurement signal to the proper form. The CNN input information in most cases is constituted by 2D [18, 19] or 3D matrix, which is due to the principle of this structure.

In the Fig. 4 the idea of processing the diagnostic signal was presented. Firstly, the voltage induced in the measuring coils was measured. The tests assumed the possibility of the damage detection after measuring 256 samples of the diagnostic signal for a sampling frequency of 2048 Hz, which corresponds to a measurement time of 0.125 seconds. Then the vectors including 256 samples of the diagnostic signals were converted to matrices of size $16 \times 16$. The last step was the conversion of two-dimensional matrices into a three-dimensional matrix. A set of two or three-dimensional arrays was then saved as two training and testing data packets.

![Fig. 4. Elaboration of the input layer of the CNN used for electrical fault detection of the induction motor – idea scheme](image)

The CNN input data were developed in two stages. The first stage consisted of the preparation of the learning and validation data applied in the training process of CNN. Based on the measurements of the diagnostic signals during the changing of the motor operating conditions and the level of damage, the packets containing 4275 and 1200 examples were developed for CNN-1 and CNN-2 respectively (Table 2). These packets constituted the basis of the training process carried out in accordance with the cross-validation technique. The next step was the development of testing data. In this case, signal measurements took place for the whole range of load torque changes, supply voltage frequency and changes in the degree of damage to the induction motor. Due to the short time of a single measurement of 0.125 s, it was decided to develop a test packet of the same size as the training packet. The implementation of two data packages enabled the analysis of the effectiveness of the developed CNN structures for data initially unknown to the convolutional neural network.

3.2. Analysis of the developed CNN structure. A characteristic feature of the convolutional networks is that they do not have a pre-established architecture, methods of parameter selection or rules regarding the number of layers. The detection of damage symptoms is related to the acquisition of generalization skills by the neural network. Due to the expanded structure of the CNN, methods to avoid over-matching become more and more important. Figure 5 presents the structure of a convolutional network consisting of 3 sets of convolution layers and a set of layers, whose task was to determine the affiliation to classes. The developed CNN structure permits to distinguish the degree of IM electrical circuits damage based on the information coming from the axial flux signal.

![Fig. 5. Developed convolutional neural network structure](image)

In the Fig. 5, the following layers of the convolutional network can be distinguished:

**Convolution layer** – acts as a feature detector using the mathematical convolution operation consisting of combining two data sets. In diagnostic applications, the convolution operation is performed for multidimensional input tables. The convolutional layers have parameters that are most often trained us-
ing a simple gradient (filters, activation maps). Convolutional layer hyperparameters are called all configuration settings invariable during network learning (filter size, input data depth, step, padding). 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 axial flux signals.

**Batch normalization layer** – normalizes the data packet mainly to speed up the training process as well as to increase the stability of neural network training. The batch normalization layer normalizes the output of the previous layer, subtracting the average value of the batch elements and dividing by a standard deviation.

**Pooling layer** – its task is to select only that piece of information whose contribution to each cell is the greatest. For this purpose, most often methods are used to search for maximum or average values from cell elements. The advantage of the pooling layers is the reduction of the spatial size of the data representation, thus preventing overmatching. The pooling layers are characterized by two hyperparameters: cell size and step. Since the pooling layers calculate the constant function of the input batch, they have no input parameters.

**Dropout layer** – is used to avoid the situation when a single neuron is strongly dependent on the state of the others. This technique allows teaching each of the neurons a different useful feature.

**Fully connected layer** – its task is to assess the share of individual classes that are the result of the network. The convolutional network contains one or more fully connected layers of any size.

**Softmax layer** – is the most frequently used activation function during the process of evaluating the output class. Its task is to determine the probability of affiliation to one of the known categories. Its task is to determine the probability of the input batch elements belonging to one of the known categories.

**Classification layer** – generates neural network response based on information whose contribution to each cell is the greatest. For this purpose, the cross-entropy of losses is calculated.

During the research, two convolutional neural structures were developed to perform the following tasks:
- **CNN-1** – classification of stator damage in phase A and rotor cage defects (9 categories).
- **CNN-2** – classification of the degree of stator damage in 3 phases of an induction motor (16 categories).

The structure parameters are listed in Table 1.

### 3.3. CNN training process.

The process of training convolutional neural networks is usually carried out according to the Stochastic Gradient Descent (SGD) method. The SGD training algorithm allows determining the gradient estimation using the average gradient of the samples drawn from the mini packets of learning data. The key parameter of the SGD algorithm is the learning rate. The value of this parameter is most often selected based on the analysis of learning curves. An alternative to the slow SGD algorithm is its extension with a momentum parameter (SGDM – Stochastic Gradient Descent with Momentum).

Momentum is a technique that allows accelerating SGD in the appropriate direction at the same time significantly reducing oscillations. The SGDM algorithm has the largest step size when multiple consecutive gradients point in the same direction. As a result, we gain faster convergence and reduced oscillation. The training process according to the SGDM algorithm starts with the initial value of learning speed $\eta$ and the momentum parameter $\alpha$. Then, for a random sample from a mini batch packet $\{x_1, \ldots, x_m\}$ the $p$-gradient is estimated according to the relationship:

$$ p = \frac{1}{m} \sum_{i} L\left( f\left( x_i, w \right), y_i \right), $$

where: $x_i$ – a randomly selected minibatch element with size $m$, $L\left( f\left( x_i, w \right), y_i \right)$ – calculated loss function for $i$-th sample.

For the estimated gradient value $p$, momentum $\nu$ and parameter $w$ are updated in accordance with the following relationships:

$$ \nu = \alpha \nu - \eta p, $$

$$ w = w + \nu, $$

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

The SGDM algorithm through the use of data packets in order to estimate the gradient approximation also allows to update the weight matrices of CNN during the experimental verification based on data from actual measurements. This is a big advantage of this technique, especially in terms of hardware im-

<table>
<thead>
<tr>
<th>Name of parameter</th>
<th>Value of parameter</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Number of convolutional layers</strong></td>
<td>CNN-1</td>
</tr>
<tr>
<td><strong>Number of filters</strong></td>
<td>32-32-64</td>
</tr>
<tr>
<td><strong>Filters size</strong></td>
<td>3.3</td>
</tr>
<tr>
<td><strong>Depth</strong></td>
<td>1</td>
</tr>
<tr>
<td><strong>Padding/Stride</strong></td>
<td>“same”/(1.1)</td>
</tr>
<tr>
<td><strong>Number of pooling layers</strong></td>
<td>3</td>
</tr>
<tr>
<td><strong>Pooling method</strong></td>
<td>maximum</td>
</tr>
<tr>
<td><strong>Pool size</strong></td>
<td>3.3</td>
</tr>
<tr>
<td><strong>Padding/Stride</strong></td>
<td>“same”/(1.1)</td>
</tr>
<tr>
<td><strong>Number of activation layers</strong></td>
<td>3</td>
</tr>
<tr>
<td><strong>Activation function</strong></td>
<td>Hyperbolic tangent</td>
</tr>
<tr>
<td><strong>Number of normalization layers</strong></td>
<td>3</td>
</tr>
<tr>
<td><strong>Value of $\varepsilon$ coefficient</strong></td>
<td>0.0001</td>
</tr>
<tr>
<td><strong>Number of dropout layers</strong></td>
<td>1</td>
</tr>
<tr>
<td><strong>Probability</strong></td>
<td>0.5</td>
</tr>
<tr>
<td><strong>Number of fully connected layers</strong></td>
<td>1</td>
</tr>
<tr>
<td><strong>Number of neurons</strong></td>
<td>9</td>
</tr>
<tr>
<td><strong>Accuracy</strong></td>
<td>99.6%</td>
</tr>
</tbody>
</table>
4. Experimental verification

4.1. Experimental setup. The experimental verification of the developed neuronal detectors of IM electrical damage was carried out on a laboratory test-bench consisting of a 1.5 kW squirrel-cage IM, powered from a frequency converter (Fig. 6). The load torque was generated by means of a DC machine connected by a rigid shaft with the tested IM (Fig. 7). The measuring system enabled the analysis of the diagnostic signal for various machine operating conditions. A schematic diagram of the measurement system is shown in Fig. 6.

The signal of the voltage induced in the measuring coils was sent via the National Instruments measurement card to the software developed in the LabVIEW programming environment.

The tested IM allowed to physically model the discussed rotor and stator defects. The change of the supply voltage frequency took place in the range \( f_s = (20\div50) \) Hz, while the adjustment of the load torque in the range \( T_L = (0\div1)T_N \). The parameters of the measuring coils used in the tests are presented in Table 4. The neural network training and testing process were carried out using the Matlab software. The tests were carried out for various types of fault:
- \( 0-5 \) shorted turns of one stator phases A, B, C,
- \( 0-3 \) damaged rotor cage bars.

The stator winding has been specially prepared so that it was possible to make a short circuit of the proper number of turns in each of the three phases of the stator. A corresponding group of coils was led out to the terminal board, and the short-circuit was physically modelled through a metallic connection. Rotor damage, in the form of broken bars, was modelled by preparing rotors with pre-drilled windings. In addition, the resulting holes were sealed to reduce the impact of the rotor imbalance.
4.2. Diagnosis of electrical damage to the induction motor.

The first stage of experimental verification concerned the CNN-1 structure. The task of the network was to assess the technical condition of the rotor and one phase of the stator. As shown in Table 2, the network input matrix was based on the measurement of voltages induced in coil number 1. The resulting data packet was a set of two-dimensional 16 × 16 tables. The network responsible for a given testing data package is shown in Fig. 8. As observed in Fig. 8, CNN’s structure is characterized by the high efficiency reaching over 99.6% for a test data package coming from 4275 measurements. Errors in the diagnostic information provided by the network resulted from similar quantitative changes in the diagnostic signal at low voltage frequency or load torque close to zero. Nevertheless, the network provides a very high level of precision whilst short signal measurement time. Another important aspect is that there is no necessity to use the measured signal pre-processing. Information contained in the diagnostic signal goes directly to the input of the convolutional network. This approach allows for a significant acceleration of the detection process as well as the limitation of the calculations to the neural network structure only.

4.3. Detection of IM stator windings short-circuit. The next stage of experimental verification concerned the CNN-2 structure. In this case, the task of the network was to assess the degree of damage as well as to determine its location (stator phase). In the first step, the influence of the shorted turns in individual phases of the stator on the root mean square values (RMS) of the voltages induced in the measuring coils were analysed (Fig. 9). The analysis of the voltage values allows concluding that the use of measuring coils placed in the 3 axes of the tested motor makes it possible both to assess the degree of
damage (Fig. 9a) and to determine the phase of its occurrence (Fig. 9b, c). Therefore, during the development of the CNN-2 structure, a three-dimensional input matrix containing information from three measurement coils was used. The network responsible for the test package was shown in Fig. 11. As observed, for the developed test set, CNN-2 provided only correct information about the technical condition of the motor. This is due to linear changes in the diagnostic signal due to an increase of the degree of stator shorted turns. The use of additional measuring coils made it possible to obtain information on changes in the magnetic field, characteristic of short circuits in individual phases.

Accordingly to the fact that the information from the coils located in the vertical and horizontal axes made it possible to obtain information about the failure phase, the impact of the input matrix configuration on network efficiency was also analysed. For this purpose, the operation of the CNN-2 structure was checked for 7 different data packages (Fig. 10). Confirmation of assumptions resulting from the analysis of the voltages values induced in individual coils (Fig. 9) is noticeable in (Fig. 10) by the high level of networks precision in which the input matrix was a configuration of the signal coming from coil number 1 with a signal coming from any of the additional measurement coils (2 or 3). The use of only one coil does not allow the determination of the phase in which the shorted turns occurred. It should be noted that the provision of additional information to the network allowed for a significant reduction in the structure (Table 1) as well as a reduction of the CNN training process. This fact results from the immediate separation by the network of the input matrix features characteristic of each damage category.

5. Conclusions

The application of deep learning neural networks presented in the article allows for a significant increase in the effectiveness of damage detection and classification systems. The discussed CNN structures are not only characterized by a high level of precision but most importantly, allowed to limit the duration of the diagnostic signal measurement to 0.125 seconds. Detection time is a priority in stator short circuit detection systems due to the nature of this type of damage. Another advantage of using DNN structures is the possibility to eliminate signal processing to isolate symptoms of damage. The approach to damage detection presented in the article allows to significantly reduce the number of necessary transformations of the diagnostic signal. Despite the difficulties in selecting the optimal structure of CNN and the parameters of the training process, the use of an appropriate diagnostic signal makes it much easier to develop an effective CNN structure. Presented use of the information contained in the flux signal made it possible both to determine the technical condition of the rotor and stator and to isolate the damaged stator phase. The discussed application of the direct connection of the measuring system with the NN structure may be an alternative to the currently used analytical diagnostic methods as well as classical neural structures.

Acknowledgements. This research was supported by the National Science Centre Poland under grant number 2017/27/B/ST7/00816.

REFERENCES


