Process of Building Artificial Neural Network for Automatic Detection of Signals from Transverse Cracks in the Rail Head

Vitalij NICHOGA[1], Ivan PRUDYUS[2], Liubomyr VASHCHYSHYN[3]

Summary
In this article the process of building artificial neural network (ANN) for automatic detection signals from transverse cracks in the rail head is described. Selection details and real signal samples (for ANN training) are also presented.

Keywords: defect, transverse crack, artificial neural network

1. Introduction

The timely detection of dangerous rail defects is extremely important since they can lead to accidents with significant material costs and human victims. Therefore, rails are systematically inspected for internal and surface defects using various non-destructive testing (NDT) techniques. The most common of which are ultrasonic and magnetic flux leakage (MFL) methods. Nowadays testing carriages based on MFL are widely used for speed inspection of the railway tracks. They allow to control state of the rail head on a depth until 7–8 mm at velocities from 20 to 80 km/h during different weather conditions.

The most important issue in all methods of NDT is selection of information about defects from defectoscopic signals. Unfortunately, at this time, experience of a wagon-defectoscope operator is the main guideline in choosing the right testing evaluation. That is why automation of the defects detection process is the basic direction for improving existing NDT facilities, implementation of which is impossible without involvement of the modern digital signal processing tools (DSP). Using of the ANN is the most appropriate solution for automatic detection of signals from defects [9].

2. Selection of the ANN type

Network type should be selected based on the complexity of the problem and the available data for training. 50 real signals from transverse crack (25 of which are presented on the fig.1) were selected for ANN training. All the signals were normalized regarding to zero level. Normalization was caused by differences in the magnetizing system settings on different wagon-defectoscopes. Generally ANN training via raw patterns does not provide quality result. Based on the characteristics of the defect signal shapes [5] was decided to pre-process the signals by continuous wavelet transform (CWT). According to the study [6] CWT is the most adaptable DSP tool for their analysis.

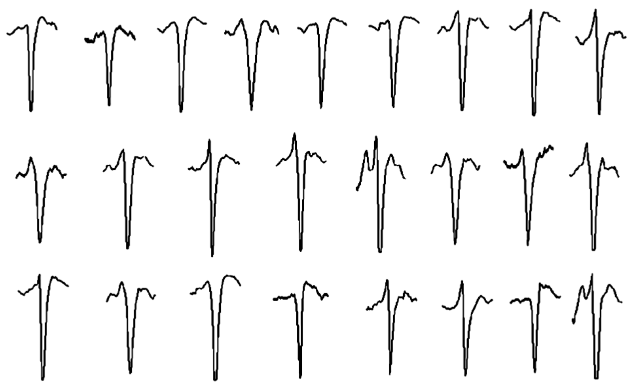

Fig. 1. 25 (from 50) real signals from transverse cracks that was used for ANN training

That is why CWT coefficients, on scales from 8 to 15, [8] for 50 signals from transverse crack (true values) and 50 signals from sleeper substrates (false values) were selected as a training set. For recognizing signals (such as signals from transverse crack) ANNs that are used for classification tasks are the most suit-

[1] Prof., DSc.; Lviv Politechnic National Uniwersity; e-mail: nich@org.lviv.net.
[2] Prof., DSc.; Lviv Politechnic National Uniwersity; e-mail: iprudyus@lp.edu.ua.
[3] Research Officer; Physico-Mechanical Institute of the National Academy of Sciences of Ukraine; e-mail: vashchyshyn.liubomyr@gmail.com.

able. Accordingly to that and available training data multilayer perceptron was selected as a type of ANN for detection defects.

3. ANN parameters selection

For proper ANN parameters selection we should follow the next rules [1]:

- the number of neurons in the hidden layer should be less than the number of elements of the training set,
- neural network performance can be increased both by number of neurons and number of layers.

The more hidden layers are present in a network the more accurate network is. However, with large dimension of hidden layer can be observed phenomenon of network retraining. In that case the network very good copes with recognizing only training set signals. That indicates the deterioration of generalizing network properties.

For multilayer perceptron it is necessary to choose the following significant parameters: number of hidden layers, number of neurons in hidden layers and the type of neurons activation function.

3.1. Determining the number of hidden layers

In [2] Timothy Masters says that to solve almost all variety of practical problems is enough only one hidden layer. Only if large number of neurons in the single hidden layer does not provide proper result should be added second. The author also stresses that theoretical necessity to use a neural network with three or more hidden layers is absent since such approach only increases training time and possibility of local minimum error during training.

That is why it was decided to use only one hidden layer for building ANN for automatic detection signals from transverse cracks in the rail head.

3.2. Determining of the optimum number of hidden layer neurons

Unfortunately, nowadays in the theory of artificial intelligence, sufficiently accurate methods for determining the number of hidden layer neurons have not been developed yet. In practice, that value should be determined empirically. However, there are heuristic rules for choosing the number of hidden layer neurons [7]. One of them is the geometric pyramid rule [2]. It states that for many practical implementations of ANN the number of neurons resembles the shape of pyramid in which quantity of neurons reduces from the input to output. According to the geometric pyramid rule number of neurons in the hidden layer of three-layered perceptron is calculated in the next way:

$$k = \sqrt{n \cdot m},$$

Where:
k – number of neurons in the hidden layer,
n – number of neurons in the input layer,
m – number of neurons in the output layer.

Applying that rule to our specific case can be calculated that for eight inputs and one output neurons the network should contain three neurons in the hidden layer. Such quantity of hidden neurons does not require a lot of computational resources.

3.3. Selecting of activation function for neurons

Method and velocity of training depend on the activation function type. Most of the activation functions have compressive properties i.e. neuron output values are always inside certain range, which depends on the type of activation function [10]. Sigmoid function (Fig. 2) is the most appropriate in our case since it provides the greatest freedom for choosing network training method and the range of its values perfectly reflects defect presence or absence (1 or 0, respectively).

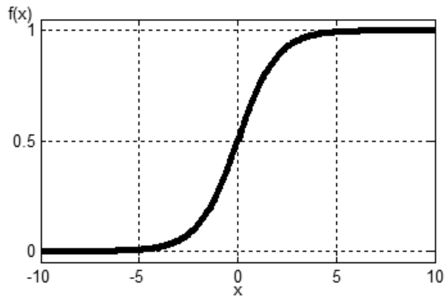

Fig. 2. Sigmoid function

4. Architecture of ANN for automatic detection of signals from transverse cracks

As a result of the aforementioned network type and parameters selection was obtained three layers ANN (Fig. 3) which consists of one input, output and hidden layers.

Functioning of the ANN can be described as follows. Input layer of the network has eight inputs on which CWT coefficients, scales from 8 to 15, [8] are passed. Wavelet adapted to detection of signals from transverse cracks [4] (Fig. 4) was selected as a mother wavelet for CWT.

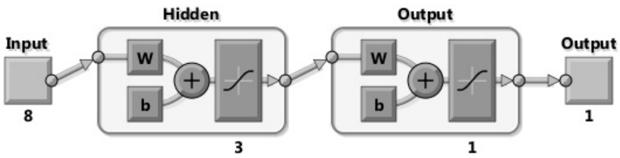

Fig. 3. ANN for automatic detection of signals from the transverse rail cracks: numbers are the amount of knots in the input, hidden and output layers, respectively

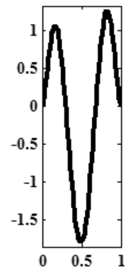

Fig. 4. Wavelet adapted to detection signals from transverse cracks in the rail head

Input signals are duplicated and fed to each of the three hidden layer nodes. They are multiplied with the weights W (which were adjusted during the network training), added and taking into account the bias node b passed to the transfer (activation) function. A bias value allows shifting the transfer function along the horizontal axis which can be crucial for successful training.

Then signals from three hidden nodes come to the output layer neuron. Depending on the similarity to the signal from a transverse crack the network generates a signal that corresponds to defect presence or absence. Results of using ANN (after training) are presented in [3] and demonstrate the ability to automatically detect signals from transverse cracks which have similar to the mother wavelet form.

5. Conclusions

1. The general process of building ANN for automatic detection signals from transverse cracks in the rail head was described.
2. The same approach can be used to construct ANN for detection other types of defects.
3. The accuracy of the ANN can be improved by the involving into processing workflow of the additional data from other NDT methods or information about previous races of testing carriage. In that case the number of input layer neurons as well as hidden layer neurons should be revised.

Literature

1. Golovko V.: *Nejronnye seti: obučenie, organizacija i priminenie*. IPRŽR, Moscow, 2001, p. 256 (in Russian).
2. Masters T.: *Practical Neural Network Recipes in C++*, Academic Press, London 1993, p. 504.
3. Nichoga V., Vashchyshyn L.: *Wavelet-neural network for detection signals from transverse cracks in the rail head* „16th International Conference on Computational Problems of Electrical Engineering (CPEE)" September 2–5, 2015, Lviv, pp. 132–134.
4. Nichoga V.O., Vashchyshyn L.V., Storozh I.V.: *Stvorennja materyns'koï vejvlet funkciï ta analiz za iï dopomogoju defektoskopičnogo sygnalu vid poperečnoï triščyny rejok,* Informacijno-kerujuči systemy na zaliznyčnomy transporti, nr 3, Charkiv 2012, pp. 61–69 (in Ukrainian).
5. Nichoga V. et al. *Doslidžennja sygnaliv magnitosdynamičnoji defektoskopiï na Lvivs'kij zaliznyci.* Physical methods and means of control environments, materials and products, vyp. 13, Lviv 2008, pp. 8–19 (in Ukrainian).
6. Nichoga V. et al. Zastosuvannja vikonnogo peretvorennja Fur'je i vejvlet peretvorennja pry analizi sygnaliv magnitnoï diagnostyky zaliznyč*nych rejok.* Proceedings of the 16th International Conference LEOTEST'2011. February 21–26, 2011 p., Slavske, Ukraine, pp. 134–139 (in Ukrainian).
7. Swingler K. *Applying Neural Networks: A Practical Guide*. Morgan Kaufman Publishers, San Francisco 1996, p. 303.
8. Vashchyshyn L.V., Nichoga V.O.: *Kryteriï vyjavlennja sygnalu vid poperečnoï triščyny rejky za dopomogoju neperervnogo vejvlet peretvorennja*. Information extraction and processing, nr 38(114), Lviv 2013, pp. 69–74 (in Ukrainian).
9. Vashchyshyn L.V., Nichoga V.O. Storozh I.V.: *Štučni nejronni mereži jak zasib dlja rozpiznavannja defektiv zaliznyčnych rejok.* Informacijno-kerujuči systemy na zaliznyčnomy transporti, nr 5, Charkiv 2012, pp. 34–37 (in Ukrainian).
10. Wasserman P.: *Neural Computing: Theory and Practice*, Van Nostrand Reinhold, New York, 1990, p. 192.

Proces projektowania sztucznej sieci neuronowej do automatycznego wykrywania sygnałów z poprzecznych pęknięć w główce szyny kolejowej

Streszczenie
Artykuł opisuje proces projektowania sztucznej sieci neuronowej (ANN) do automatycznego wykrywania sygnałów z pęknięć poprzecznych w główce szyny kolejowej. Do celów szkoleniowych w zakresie ANN przedstawiono również szczegóły dotyczące wyboru próbki i rzeczywiste próbki sygnału.

Słowa kluczowe: uszkodzenie, pęknięcie poprzeczne, sztuczna sieć neuronowa

Процесс проектирования искусственной нейронной сети для автоматического обнаружения сигналов из поперечных трещин головки железнодорожного рельса

Резюме
В статье описан процесс проектирования искусственной нейронной сети (ANN) для автоматического обнаружения сигналов из поперечных трещин головки железнодорожного рельса. Представлены также подробности выбора и действительные образцы сигнала (для обучения ANN).

Ключевые слова: повреждение, поперечная трещина, искусственная нейронная сеть