International Journal of Applied Engineering Research ISSN 0973-4562 Volume 12, Number 12 (2017) pp. 3461-3469

Acoustic Signals Recognition by Convolutional Neural Network

Vera Barat[1], Peter Kostenko[1], Vladimir Bardakov[1] and Denis Terentyev[2],

[1]Department of Electrical Engineering & Introscopy,

National Research University "Moscow Power Engineering Institute", Moscow, Russia.

[2]INTERUNIS-IT LLC, 111024 Moscow, Russia.

Abstract

Acoustic emission is non-destructive method which is widely and successfully used to assess the technical condition of industrial facilities. It is challenging to divide useful signals and a noise in the flow of acoustic emission data. In this paper we describe the application of convolutional neural networks to classify acoustic emission signals by their time-frequency structures. Convolutional neural network was learned to classify acoustic emission signals with 95% accuracy. Acoustic signal recognition by convolutional neural networks can greatly simplify the data processing.

Keywords: acoustic emission; time-frequency analysis; wavelet; convolutional neural networks.

INTRODUCTION

The problem of technical and environmental security is essential for modern society. The complexity and variety of building structures, high intensity level of industrial production necessitate an expert examination of industrial safety and non-destructive testing of industrial installations, buildings and structures.

One of the methods of nondestructive testing is based on the phenomenon of acoustic emission (AE) [1], which consists in the generation of elastic waves under structural changes in the material - deformations, phase transformations, formation of defects. The transducer installed at the monitoring facility records acoustic emission waves emitted by the defect and transforms them into electrical signals. The parameters of the measured signals allow to determine the location of the defect and the degree of its danger.

At present, the AE method is widely and successfully used to assess the technical condition of industrial facilities. It is remote method and allows during a short time, in comparison with other methods, to control the entire object, the size of which can reach several tens of meters. By its resolution ability to the cracking process, it significantly exceeds the capabilities of traditional methods of nondestructive testing and makes it possible to detect a crack-like defect even at the stage of nucleation.

The disadvantage of AE control is the sensitivity to the noise accompanying the testing procedure. AE control often results in the registration of a large data stream containing both useful signals caused by growth of a defect and noise - electromagnetic interference, external acoustic noises, noise associated with friction, and also with the technological processes of the monitored object.

The specificity and complexity of noise filtering during AE control lie in the fact, that, as a rule, noise and useful signals have the same frequency range, so that signal recognition and classification methods are used to isolate useful signals from interference. The classifier is constructed in such a way that the signals characterizing the AE activity of defects are referred to one particular class, and the noises are referred to another or several other classes.

In this paper, the use of a convolutional neural network is considered to classify the flow of AE data. An image - a spectrogram of the AE signal [2-4] - is sent to the input of the

International Journal of Applied Engineering Research ISSN 0973-4562 Volume 12, Number 12 (2017) pp. 3461-3469
 http://www.ripublication.com

neural network. The neural network allocates AE signals corresponding to defects, according to the presence of a characteristic time-frequency structure of the signal.

CHARACTERISTICS OF SIGNALS AND NOISE

The process of the AE signal forming is shown in Figure 1. Defect 1, present in the material, emits AE waves 2. Propagating along the acoustic path of the monitoring object, AE waves reach the primary transducer (TAE) 3, which transforms the acoustic signal into an electric signal.

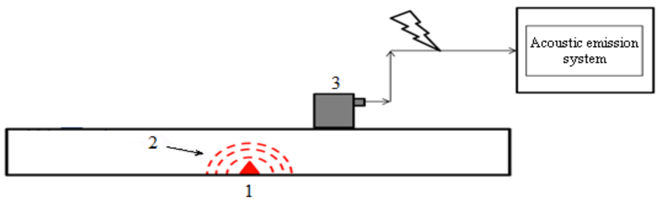
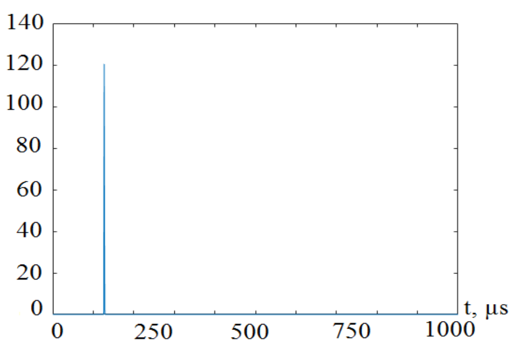


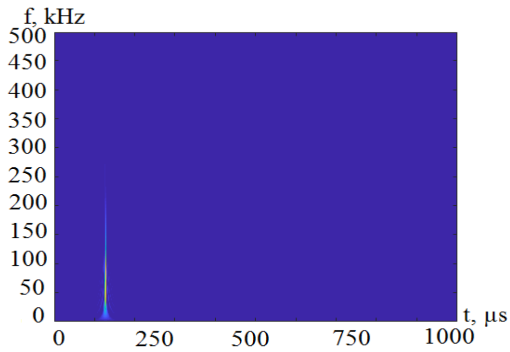
Figure 1. The process of AE signal forming. 1 – the source of acoustic emission, 2 – AE waves, 3 – primary transducer.

The AE signal, which is registered by the primary transducer, can be represented, in some approximation, as a result of the convolution of the source function and the acoustic transfer function.

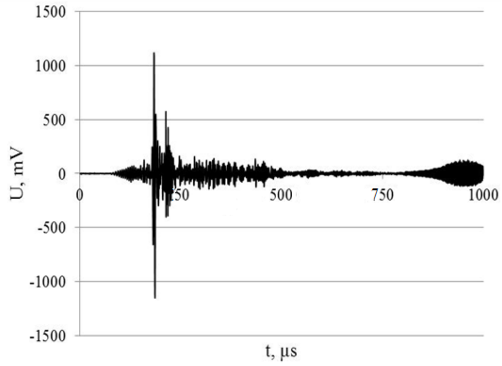
If the source of acoustic emission is a defect, the function $u_{source}(t)$ is a short pulse of duration on the order of fractions of a microsecond, which can be represented, in some approximation, as a delta function (Fig.2a).



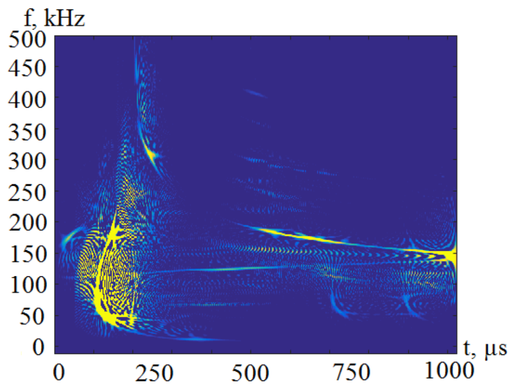
(a)



(b)



(c)



(d)

Figure 2. (**a**) Delta function in time domain (**b**) Delta function in frequency domain (**c**) Signal of acoustic emission in time domain (**d**) Signal of acoustic emission in frequency domain.

The impulse response of the acoustic path is more complex. In most cases AE control is carried out for thin-walled objects with a wall thickness of 3-30 mm. The AE wavelength is comparable to the thickness of the object wall, and the AE waves propagate in the form of normal (waveguide) waves, usually Lamb waves [5]. Such waves differ in dispersion - the dependence of the propagation velocity on frequency. The source of the AE simultaneously at a time T emits a short broadband wave packet (Figure 2 (**a**), (**b**)), in accordance with the wave equations for the case of Lamb waves [4], the wave packet components corresponding to different frequencies propagate at different rates. The times they are registered are given by the relation (1)

$$t(f)=\begin{cases} t_{S_0}(f)=T+L/V_{S_0}(f,h,V_L,V_T) \\ t_{A_0}(f)=T+L/V_{A_0}(f,h,V_L,V_T) \end{cases} \quad (1)$$

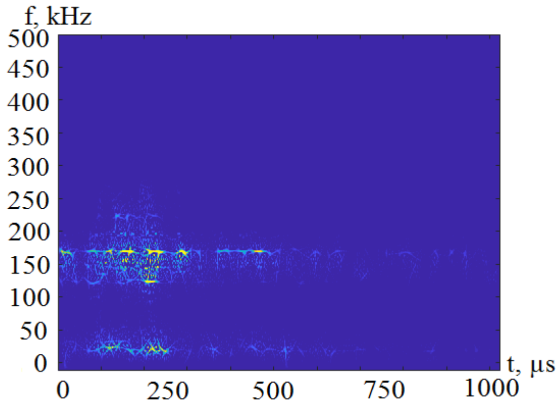
Here V_{S_0} and V_{A_0} are the group velocities of the two main Lamb waves S_0 and A_0 [5], V_L and V_T are the velocities of the bulk longitudinal and transverse waves in the object material.

The dispersion pattern of AE signal propagation is manifested during time-frequency analysis of the signal in the form of curves corresponding to the shape of the dispersion curves of Lamb waves. In this paper, the Wigner-Ville transform is used as the time-frequency transformation, which allows to analyze the energy distribution of the signal with respect to time and frequency, [7, 9].

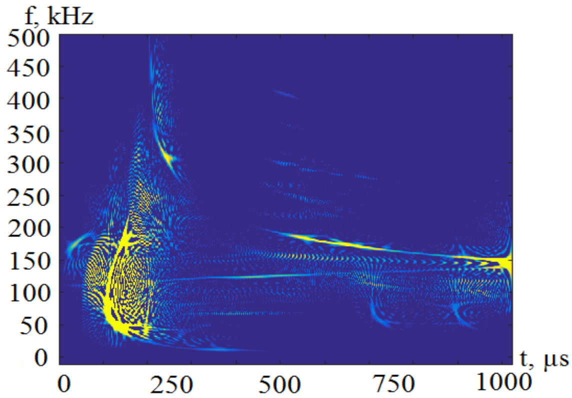
Figures **2**(**b**) and **2**(**d**) show the Wigner transformation (spectrogram) for the delta function of the source $u_{source}(t)$, and for the signal $s(t)$ that passed through the acoustic path of the monitoring object. On the spectrogram of the signal $u_{source}(t)$, the signal energy is distributed over a wide range of frequencies, but is localized in time. The spectrogram of the $s(t)$ signal is the one of dispersive propagation of the AE signal. At the time-frequency plane, there are curves characterizing the dispersive propagation of symmetric S_0 and asymmetric A_0 Lamb waves [5].

Noises arising during AE testing are very diverse [8], some of them are electrical and electromagnetic noise, as well as background stationary noise, which are suppressed by frequency filtering. The greatest complexity for filtering is the acoustic noise - friction noise, various shock effects, in particular the influence of precipitation. Acoustic noise in the spectrum and shape are similar to AE signals. However, with time-frequency analysis of noise signals, the characteristic time-frequency pattern is not detected.

This is explained by the fact that the function of the u_{source} for noise has a substantially longer duration, of the order of 100 μs, when such a function is convolved with the transfer function of the path, the resulting function will not have a clear dispersion structure.



(a)



(b)

Figure 3. (**a**) – The acoustic signal of rain drop in frequency domain (**b**) – a acoustic signal caused by the destruction of the material.

Figure 3(**a**) shows the acoustic signal, the source of which is a drop of rain, in Figure 3(**b**) - a signal caused by the destruction of the material. Both signals have a pulsed non-stationary

International Journal of Applied Engineering Research ISSN 0973-4562 Volume 12, Number 12 (2017) pp. 3461-3469

character, the spectra of both signals are broadband, non-uniform with a maximum corresponding to the resonant frequency of the transducer. Thus, only the presence or absence of dispersion curves on the spectrogram can be a reliable sign for the classification of signals and noise.

However, the detection of dispersion curves is quite a complicated task, since the spatial position of the crack and its type, the signal attenuation, the amplitude-frequency characteristic of the TAE, and also the reflection of the waves from the edges of the object, influence the shape of the AE signal and, therefore, its spectrogram form, which leads to distortion of the dispersion pattern of the AE signal and substantially complicates its form. In addition, as can be seen from Figure 4, in most cases only a small section of the dispersion curves can be seen on the spectrograms.

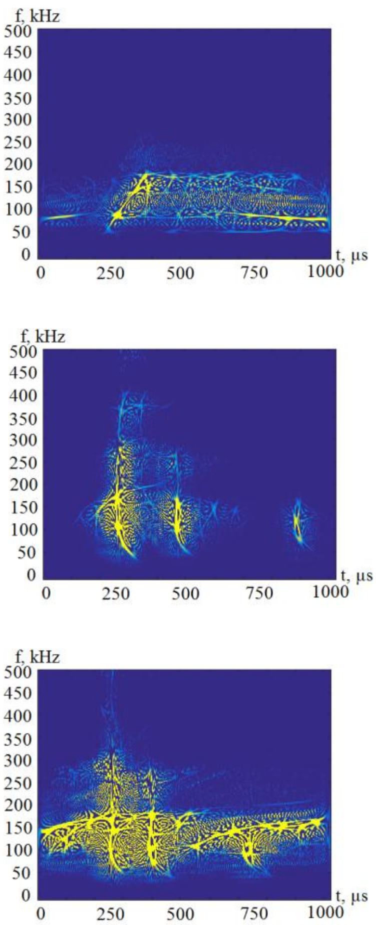


Figure 4. Examples of acoustic signals caused by the destruction of material in frequency domain.

In view of the large volumes of AE control data and the absence of clear formalized features on which automatic recognition could be carried out, an artificial neural network modeling the cognitive abilities of the intellect inherent in man was used to identify the dispersion pattern in this paper.

CLASSIFICATION OF SPECTROGRAMS USING AN ARTIFICIAL NEURAL NETWORK

The use of a convolutional neural network - a special architecture aimed at efficient image recognition [10] – is the one of most suitable to solve this problem.

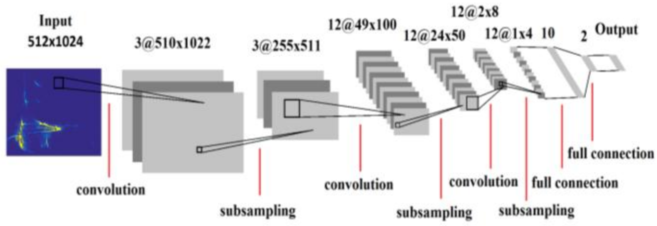


Figure 5. Main layers of the used network

A convolutional neural network is resistant to shear, scaling, noisiness, that is, to all types of transformations to which the dispersion curves (2) of Lamb waves are subjected on spectrograms.

A typical structure of a convolutional neural network is shown in Fig.5. It consists of two-dimensional layers of neurons, through which the input image is recognized. A convolutional neural network performs two main functions - image recognition or highlighting of its characteristic features, and classification based on the training sample.

Image recognition:

The selection of image features is carried out in the layers of convolution and subsample, which are the main layers of the neural network. In the convolutional layers, feature maps are formed. The feature map is the result of the convolution (2) of the input values of neurons with kernels coefficients $w_{p,s}^{(n,k)}$, which are parameters of this layer.

$$x_{i,j}^{(n,k)} = \varphi_n \left(b^{(n,k)} + \right.$$

$$\sum_{p=1}^{P^{(n)}} \sum_{s=1}^{S^{(n)}} w_{p,s}^{(n,k)} x_{i+g^{(n)}(p-1),j+h^{(n)}(s-1)}^{(n-1)} \Big); \tag{2}$$

$$\begin{cases} i = \overline{1, I^{(n)}};\ j = \overline{1, J^{(n)}} \\ \quad k = \overline{1, K^{(n)}}; \end{cases} \tag{3}$$

$x_{i,j}^{(n,k)}$– output values of neurons in layer n, k - number of convolution kernel, $P^{(n)} \times S^{(n)}$ - sizes of convolution kernels, $g^{(n)}$ and $h^{(n)}$ - convolution step, - neuron parameter (displacement).

Each kernel $w_{p,s}^{(n,k)}$ is a description of the primitive element (feature) inherent in the image. The convolution kernel can be learned to descript such features as a line, an arc, or any other fragment of the image. The local maxima of the feature map correspond to the location on the input image of those features that are specified by the convolution kernel. For each convolution layer a set of convolution kernels is used. The convolution kernels are formed while network learning.

The convolution layer activation function is a ReLU (Rectified Linear Units), which is an extremely useful component of convolutional neural network and makes it possible to significantly speed up network training and it also helps to alleviate the vanishing gradient problem, which is the issue where lower layers of the network train slowly because the gradient decreases exponentially through the layers [11, 15].

After each convolution layer, a subsample layer follows. On the subsampling (pooling) layers, the size of the feature cards is reduced, that is called subsampling (pooling). Subsampling allows to speed up the learning process of the network, as well as to perform a search for features at a larger scale. Subsampling is carried out according to a fragment of the features map by replacing its elements with an average or maximum (most common) value or using another function.

Convolutional neural networks are part of the deep learning technology. The deep learning function is implemented by repeating layers of convolution and subsampling, which allows to model complex objects that have a high level of abstraction by building a hierarchy of simpler abstractions. In repeated layers the number of feature maps increase, and their sizes decrease, until the two-dimensional feature map is transformed into a one-dimensional feature vector containing the description of the original image.

Image classification:

The final element of the convolutional neural network is a one-dimensional, fully connected neural network similar to a multilayer perceptron. The input values of this neural network are feature vectors that were formed on the last pooling layer. The values of the neurons of the output layer determine the probability that an input image belongs to one or another class. The activation function of a fully connected neural network can be a sigmoidal function or a hyperbolic tangent, or any other activation function used in constructing a multilayer perceptron.

Training of a convolutional neural network:

Network training is conducted through training with the teacher, when a set of training signals $X_q^{(1)} = \left\{x_{i,j}^{(1)}\right\}_q$ is sent to the input of the network, for each of which the correct output is known (the desired response) C_q (pair {1;0} or {0;1}), in the process of such training network parameters such as convolutional kernels, synaptic coefficients of a fully connected neural network are adjusted to achieve the best match between the desired and real output, N is the number of the output layer. Training is conducted by the method of stochastic mini-batch gradient descent [12].

The function of cross-entropy is used as a function of loss, characterizing the difference between the desired and the real result at the output of the network [13]:

$$E = -\frac{1}{k} \sum_{q=1}^{Q} \ln p_q, \tag{5}$$

Where Q - the number of training signals, p_q - the probability of the belonging of the training example $\left\{x_{i,j}^{(1)}\right\}_q$ to its class C_q given by the classifier network:

$$p_q = \begin{cases} \left\{x_1^{(N)}\right\}_q, if \left\{x_{i,j}^{(1)}\right\}_q \ corresponds\ to\ class\ \text{№}1 \\ \left\{x_2^{(N)}\right\}_q, if \left\{x_{i,j}^{(1)}\right\}_q \ corresponds\ to\ class\ \text{№}2 \end{cases}, \tag{6}$$

Convolutional neural network for dispersion curves recognition:

A convolutional neural network was used to recognize dispersion curves on the spectrograms of acoustic emission signals (Figure 5). After considering several network configurations, the network architecture was selected to provide the required recognition accuracy. The architecture of the neural network is shown in Table 1.

The network contains three repeating pairs of convolution and subsampling layers. In the first convolutional layer, the small size (3x3) kernel was chosen to filter the Wigner-Ville transform artifacts, the two subsequent convolution layers were used to isolate large-scale features and describe the image. The activation function is the ReLU (Rectified-Linear Unit).

$$\varphi_n(x) = \max(0, x) = \begin{cases} x, x > 0 \\ 0, x \leq 0 \end{cases}; \quad (4)$$

A fully connected neural network consists of two layers, the output layer contains two neurons that correspond to two classes of images – one containing dispersion curves and one that do not.

Table 1: Architecture of the network

№	Layer type	Layer functions	Number of features	Size
1.	input		1	512x1024
2a.	convolution	Number of kernels 3, kernel size 3x3, convolution step 1	3	510x1022
2b.	subsample	Subsampling with average, matrix size 2x2, matrix step 2	3	255x511
3a.	convolution	Number of kernels 12, kernel size 15x15, convolution step 5	12	49x100
3b.	subsample	Subsampling with maximum, matrix size 2x2, matrix step 2	12	24x50
4a.	convolution	Number of kernels 12, kernel size 15x15, convolution step 5	12	2x8
4b.	subsample	Subsampling with maximum, matrix size 2x2, array step 2	12	1x4
5.	fully connected	Number of neurons	10	10x1
6.	fully connected	Number of neurons	2	2x1
7.	output	-		2x1

Training and test data sets :

The training data set consisted of 3000 spectrograms of AE signals, 1500 of which contained dispersion curves, and 1500 did not contain them.

(a)

(b)

(c)

(d)

Figure 6. Objects of control, data from which was used in data set. (**a**) A spherical gas-holder with a wall thickness of 39 mm (**b**) Jumper of the gas pipeline with a wall thickness of 17 mm (**c**) Arrow of a career dragline with a wall thickness from 9 mm to 10 mm (**d**) Part of a pipeline with changing wall thickness from 6 mm to 12 mm.

For the first group we used AE signals emitted by real sources of AE defects, as well as electronic and mechanical defect imitators located at distances up to 30 m from the TAE. Pipelines with a wall thickness of 6 to 17 mm and a spherical gas-holder with a wall thickness of 39 mm were used as testing objects. The second group of signals was formed by different noises, arising while controlling the objects described above. The sample included electrical and electromagnetic pickups, friction noise, noise from rain and hail, and mechanical noise induced by operation of the monitored object.

The test data set consisted of 1500 spectrograms, half of which were obtained on the above objects, but did not participate in the training. The second half of the test data set was generated by the data of the AE control of dragline bearing structures; these data were not included in the training sample.

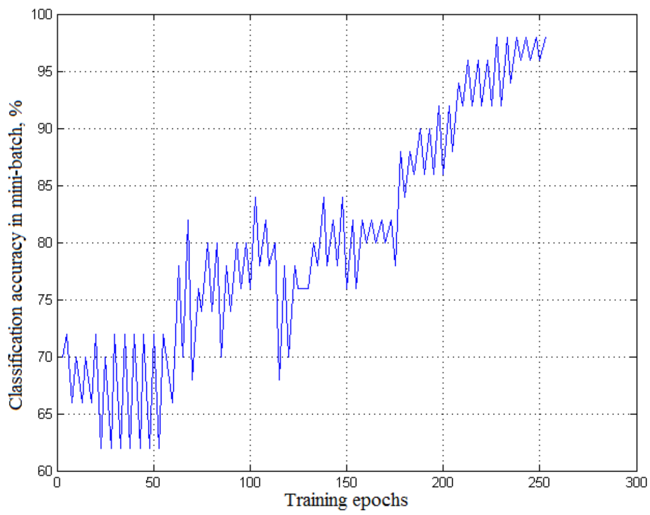


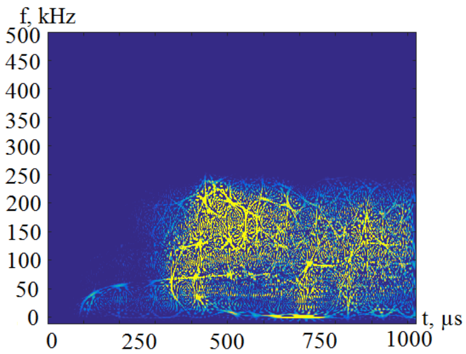
Figure 7. Schedule of network training accuracy for a mini-batch

During the 253 editions of training, the classification accuracy on the training data set was 98% (Figure 7), With further training of the network, the effect of retraining occurs, the classification error for the training data set decreases, and for the test data set increases sharply.

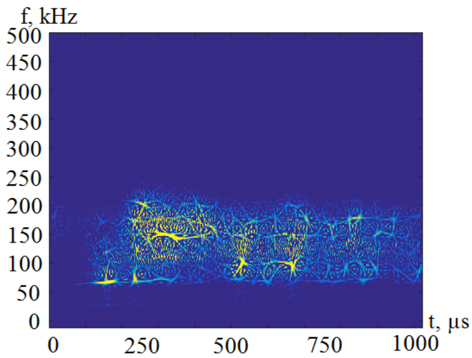
The neural network classification matrix for the test sample is shown in Table 2. Examples of spectrograms recognition are shown in Figure 8.

Table 2. The classification results

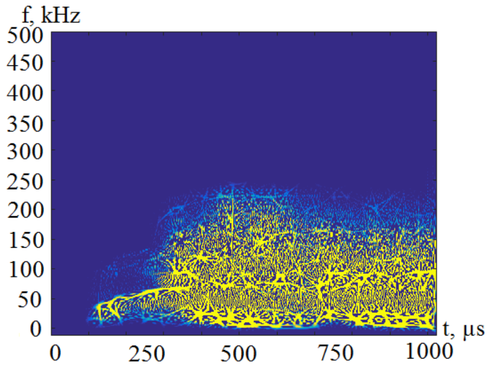
	Correct Classification	Wrong classification
Spectrograms with dispersion curves	98,6%	1,4%
Spectrograms without dispersion curves	93,5%	6,5%



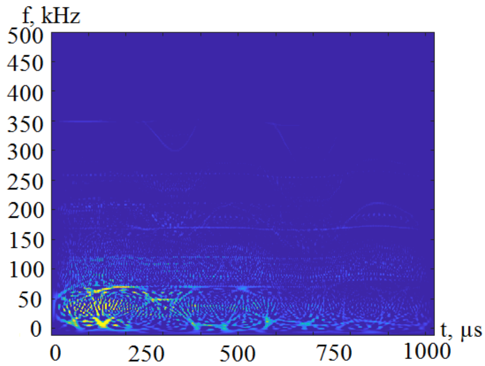
The probability of dispersion curves presence 91,72%



The probability of dispersion curves presence 4,26%



The probability of dispersion curves presence 88,17%



The probability of dispersion curves presence 6,08%

Figure 8. Examples of spectrograms recognition

CONCLUSIONS

Convolutional neural networks allow to isolate AE signals from noise by recognizing the dispersion structure on the spectrogram of the AE signal. The accuracy of the AE signals classification reaches 98% on the training data set and 95% on the test data set.

Further research in this direction can greatly simplify the process of acoustic emission data processing and increase the reliability of monitoring results.

ACKNOWLEDGMENTS

The results obtained during the execution of the State Assignment of the Ministry of Education and Science of the Russian Federation in the field of scientific activity (Project Code 11.9879.2017/8.9)

REFERENCES

[1] Rules of organization and carriage of acoustic-emission testing of vessels, apparatus, boilers and technical pipelines (approved by decree of

Gosgortechnadzor RF 09.06.2003 №77). P. 5.5. Determination of location of acoustic emission sources.

[2] Hamstad M.A.; O'Gallagher A., Gary J. Examination of the Application of a Wavelet Transform to Acoustic Emission Signals. – Journal of Acoustic Emission, 2002, v. 20, p. 39-81.

[3] Bykov S. P.; Yushin A. V.; Morozov I. L. Features of acoustic emission control of bridge cranes. – Abstracts of the conference "Risk management. Industrial safety, control and monitoring". Scientific and industrial union "RISKOM", 3-6 October 2006. http://himmash.irk.ru/ru/content/view/194/24/

[4] Terentyev D.A.; Elizarov S.V. Wavelet analysis of AE signals in thin-walled objects. – Testing. Diagnostics, 2008, №7, p. 51-51.

[5] Victorov I.A. The physical bases for the application of Rayleigh and Lamb ultrasonic waves in engineering. M.:Science. 1966. – p. 169.

[6] Hamstad M.A. Acoustic Emission Signals Generated by Monopole (Pencil Lead Break) Versus Dipole Sources: Finite Element Modeling and Experiments. – Journal of Acoustic Emission, 2007, v. 25, p. 92–106.

[7] Terentyev D.A. Identification of acoustic emission signals using frequency-time analysis. – NDT World. 2013. №2. p. 51-55.

[8] Terentyev D.A.; Bulygin K.A. New methods of applying normal waves for the control of thin-walled objects of large geometric dimensions. Part 1: Automatic recognition of dispersion curves on the spectrogram of the AE signal. – NDT World, 2011. №2. p. 46-48.

[9] Terentyev D.A.; Popkov U.S. Determination of parameters of the dispersion curves of Lamb waves by means of the Huff transformation of the AE signal spectrogram. – Defectoscopy. 2014. №1. p. 25-36.

[10] LeCun, Yann; Léon Bottou; Yoshua Bengio; Patrick Haffner (1998). "Gradient-based learning applied to document recognition" (PDF). Proceedings of the IEEE. 86 (11): 2278–2324. doi:10.1109/5.726791.

[11] Nair, V.; Hinton G.E. Rectified linear units improve restricted boltzmann machines. In Proc. 27th International Conference on Machine Learning, 2010.

[12] Qi Qian; Rong Jin; Jinfeng Yi; Lijun Zhang; Shenghuo Zhu. Efficient Distance Metric Learning by Adaptive Sampling and Mini-Batch Stochastic Gradient Descent (SGD). Department of Computer Science and Engineering Michigan State University.

[13] Michael Nielsen "Neural Networks и Deep Learning". http://neuralnetworksanddeeplearning.com/

[14] Adit Deshpande. A beginner's guide to understanding convolutional neural networks. https://adeshpande3.github.io/A-Beginner's-Guide-To-Understanding-Convolutional-Neural-Networks/