НЕЙРОІНФОРМАТИКА ТА ІНТЕЛЕКТУАЛЬНІ СИСТЕМИ

NEUROINFORMATICS AND INTELLIGENT SYSTEMS

UDC 004.93

LIGHTWEIGHT MULTI-SCALE CONVOLUTIONAL TRANSFORMER FOR AIRCRAFT FAULT DIAGNOSIS USING VIBRATION ANALYSIS

Didenko Andrii Y. – Postgraduate student of the Department of Software Tools, National University "Zaporizhzhia Polytechnic", Zaporizhzhia, Ukraine.

Didenko Artem Y. – Postgraduate student of the Department of Software Tools, National University "Zaporizhzhia Polytechnic", Zaporizhzhia, Ukraine.

Subbotin S. A. – Dr. Sc., Professor, Head of the Department of Software Tools, National University "Zaporizhzhia Polytechnic", Zaporizhzhia, Ukraine.

ABSTRACT

Context. Fault diagnosis in rotating machinery, especially in aircraft, plays an important role in health monitoring systems. Early and accurate fault detection can significantly reduce the cost of repair and increase the lifetime of the mechanism. To detect the fault efficiently, intelligent methods based on traditional machine learning and deep learning techniques are used. The object of the research is the process of detecting faults in aircraft based on vibration analysis.

Objective of the work is the development of a deep learning method for fault diagnosis in rotating machinery with a high accuracy rate.

Method. The proposed method employs Transformer architecture. The first stage of processing the vibration signal is the multi-scale feature extractor. This stage allows the model to examine input signals in different scales and reduce the impact of the noise. The second stage is the Convolutional Transformer neural network. The convolution was introduced to the Transformer to combine locality and long-range dependencies feature extraction. The Self-attention mechanism of the Transformer was changed to Channel Attention, which reduces the number of parameters but maintains the strength of the attention. To maintain this idea, similar changes were made in the position-wise feed-forward network.

Results. The proposed method is tested on the aircraft vibration dataset. Two conditions were chosen for testing: limited data and noisy environment. The limited data condition is simulated by selecting a small number of samples into the training set (a maximum of 10 per class). The noisy environment condition is simulated by adding Gaussian noise to the raw signal. According to the obtained results, the proposed method achieves a high average precision metric rate with a small number of parameters. The experiments also show the importance of the proposed modules and changes, confirming the assumptions about the process of feature extraction.

Conclusion. The results of the conducted experiments show that the proposed model can detect faults with almost perfect accuracy, even with a small number of parameters. The proposed lightweight model is robust in limited data conditions and noisy environment conditions. The prospects for further research are the development of fast and accurate neural networks for fault diagnosis and the development of limited data training techniques.

KEYWORDS: fault detection, deep learning, rotating machinery, signal processing, transformer, neural networks.

ABBREVIATIONS

AP is an Average Precision;
CA is a Channel Attention;
CNN is a Convolutional Neural Network;
DL is a Deep Learning;
FD is a Fault Detection;
FFN is a Feed-forward Network;
HUMS is a Health and Usage Monitoring System;
LMCT is a Lightweight Multi-scale Convolutional Transformer;
MHSA is a Multi-head Self-attention;
ML is a Machine Learning;
MLP is a Multi-layer Perceptron;
MSFE is a Multi-scale Feature Extractor;
NLP is a Natural Language Processing;
RNN is a Recurrent Neural Network;
SNR is a Signal-to-noise Ratio.

NOMENCLATURE

T is a time-series sensor output;
t_i is a datapoint of sensor output;
Y is a class label;
D is a data set;
M is a model;
J is a classification metric;
P_s is a signal power;

DOI 10.15588/1607-3274-2025-1-7

P_n is a noise power;

N_p is a number of patches;

f_{in} is a set of input features to the network module;

f_{out} is a set of output features of the network module;

f_{avg} is a set of features from the average pooling layer in CA module;

$f_{\max}$ is a set of features from the max pooling layer in CA module;

f^* is a set of features from intermediate layers in network modules.

INTRODUCTION

Rotating machinery is an integral part of mechanical systems. It finds its applications in many industrial sectors, including aircraft, wind turbines, pumps, car engines, etc. In aircraft and helicopters, rotating components like bearings, rotors, and gearboxes are critical for efficient operations. HUMS are sensor-based systems designed to monitor the health and performance of critical components as they are often subject to potential failure due to continuous mechanical stress (overload, overheating, lack of lubrication, etc). The early FD of these parts ensures the reliability and safety of the vehicle's exploitation, reducing the risk of critical failures and the repair cost.

Vibration-based analysis is a common FD technique in rotating machinery. Traditional ML methods often employ hand-crafted features combined with statistical-based algorithms like SVM [1], kNN [2], or decision trees [3]. While these methods can be used for fault prediction, they still struggle to capture complex features and require the complicated process of manual feature engineering. Recently, DL methods have gained popularity in the field of vibration analysis in FD [4], showing promising results in solving this task. The main advantage is the ability of DL-based models to capture complex data patterns and learn features of the input data, which eliminates the need for extensive feature engineering. (MLP [5] are commonly used in classification, CNNs [6–8] and RNNs [9, 10]. Also, a combination of CNN and RNN can be used [11, 12] to better analyze both local and temporal features of the sensor data. The recent popularity of the Transformer architecture [13] in solving sequence-based problems (including sensor data) stimulates the development of the Transformer-based FD methods.

Despite the exceptional performance of the DL-based methods for FD, these methods often need a large number of layers and parameters to extract comprehensive features, which can significantly increase the size of the model, training, and inference time and make the usage inefficient under limited resource conditions.

Moreover, one of the main issues with fault diagnosis is the limited amount of data. Usually, the number of positive samples (containing the fault) in the dataset is much smaller than the number of negative samples (healthy condition). This imbalance can significantly reduce the accuracy of the classification model, making predictions unreliable. To address this issue, techniques like zero-shot [14], few-shot [15], or transfer training can be used.

The object of the research is the process of detecting faults in aircraft health monitoring systems.

The subject of the research is the deep learning method to detect faults using vibration-based analysis.

The purpose of the research is to develop and evaluate an efficient deep learning model to classify sensor data from aircraft health monitoring systems.

1 PROBLEM STATEMENT

Let $T = \{t_i \mid i = 1\ldots n\}$ be a vector of real values t_i of a time-series sensor data. A class label $Y \in \{0,1\}$ represents either an absence or a presence of a fault. Let $D = \{(T_i, Y_i) \mid i = 1\ldots N\}$ be a labeled dataset that contains ordered pairs (T_i, Y_i). The objective is to choose such model architecture M that reaches the maximum classification accuracy, i.e., $J(Y_i, M(T_i)) \rightarrow \max$ for each $(T_i, Y_i) \in D$.

2 REVIEW OF THE LITERATURE

DL methods are widely used in rotating machinery FD tasks and have proven advantages over traditional ML methods [16]. Among them, CNNs play a significant role. Due to their strong feature extraction capabilities and local receptive field, CNNs can be successfully applied to solve vibration-based classification. As a vibration signal is one-dimensional data, 1D CNNs are the common choice. Vibration signals or other sensor data often contain localized patterns that indicate anomalies and faults, and by sliding 1D filters across the time-series data, 1D CNNs can effectively capture local, time-dependent changes that may signal an occuring fault. WDCNN method [17], for example, uses 1D CNN with wide kernels in convolutional layers, which helps to reduce noise and extract features from the input signal. DSNR method [18] employs deep residual architecture with soft thresholding (shrinkable function) to suppress redundant features and reduce the effect of the noise. Zhang et al. [19] also use a deep residual 1D CNN model to extract local data features to analyze faults. Here, residual learning helps to design deeper architectures to extract more complicated features. The authors of the AMMFN method [20] proposed the use of multi-sensor input data (vibration and current) processed by the 1D CNN model with the special attention-based fusion module that helps to extract features at different hierarchical levels and correlation information between sensors' signals.

Some methods combine 1D CNN with RNNs [11, 12]. This synthesis can improve the overall model performance by leveraging the strengths of each model type. This hybrid approach allows the network to capture both local patterns (via the CNN) and long-term dependencies (via the RNN) in time-series data. While these methods can achieve promising results, they require more training time due to the sequential nature of the RNNs.

Two-dimensional CNNs can also be used for sensor data analysis. To do this, the 1D signal is converted from the time domain to the time-frequency domain, and the obtained 2D spectrogram is used for further processing by the network. Verstraete et al. [21] proposed the use of 2D

DOI 10.15588/1607-3274-2025-1-7

CNN to analyze the Short-time Fourier Transform representation of the vibration signal. Pham et al. [22] employed the VGG16 model to analyze faults under inconsistent working conditions, showing excellent results. Although 2D CNN shows good performance, they require more samples to train in order to defeat overfitting. Moreover, 2D CNN usually requires more layers to extract deep features, which increases the model's size and inference time.

In recent years, Transformer-based models have become popular in many fields, including NLP [13], Computer Vision [23], Time-series analysis [24], FD [25], etc., outperforming common deep learning techniques. The main idea of Transformers lies behind the attention mechanism that allows focusing on the most relevant features and learning long-range dependencies more effectively. For example, the TST method [26] uses a 1D Transformer to directly analyze raw vibration signal data without preprocessing, showing high accuracy. The other studies [27] use 2D Transformers, called Vision Transformers [23], to analyze the spectrogram representation of the signal time-frequency domain. For example, ECTN [28] combines CNN and 2D Transformer to efficiently extract local and global information of the input signal. Integrated ViT model [29] decomposes the input signal with a Discrete Wavelet Transform and then applies soft voting to combine preliminary results.

Despite the advantages of the Transformers, they also have several notable disadvantages, including quadratic complexity of the attention mechanism and lack of intrinsic inductive bias [30], which makes them difficult to train. To address this issue in FD, the convolution can be introduced to Transformers [28, 31]. TWC method [32] is used in aircraft engine bearing FD and applies convolutional feature extraction layers to obtain features from raw input signal before passing it to the Transformer backbone. In TCN [33], a combined model of convolution and Transformer is trained with transfer learning technique.

Usually, the changes in vibration signal that refer to faults can vary in scale. Thus, using single-scale signal analysis, the fault can be easily overlooked. Multi-scale analysis can be used with different DL models. MCF-1DViT [25] uses a Multi-scale Convolution Fusion Layer to extract features at different time scales and pass them to the Transformer. AM-CNN [34] employs a Multi-scale convolutional block with CNN for the same purpose. Chen et al. [35] proposed MCNN-LSTM model with low-frequency and high-frequency feature extraction branches. Saghi et al. [11] developed CNN with three-scale branches followed by a bidirectional GRU model. These studies show that multi-scale feature extraction can effectively capture more information at different scales and enhance the overall performance of the method.

Recent studies demonstrate the potential of DL-based methods for FD in rotating machinery. However, further improvements are needed to reduce model complexity and the amount of noise in the input signal, achieve scale invariance, and improve local and global feature extraction, emphasizing the relevance of the proposed method.

DOI 10.15588/1607-3274-2025-1-7

3 MATERIALS AND METHODS

To solve the above challenges, the novel FD method called LMCT was developed. The proposed method mainly consists of two parts: MSFE and Transformer encoder. Fig 1. depicts the overall architecture of the neural network.

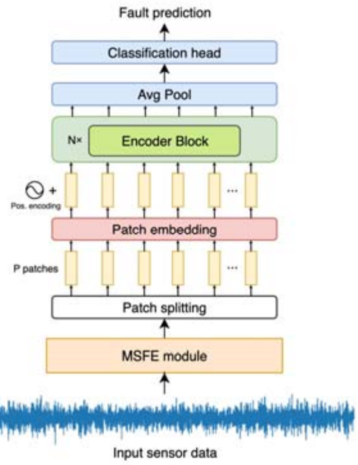

Figure 1 – Architecture of the proposed method

Both MSFE and Transformer encoder employ CA [36]. CA allows model to adjust the importance of each channel, by helping it to focus more on the informative ones while paying less attention to those that contribute less. The architecture of the CA module is show in Fig 2.

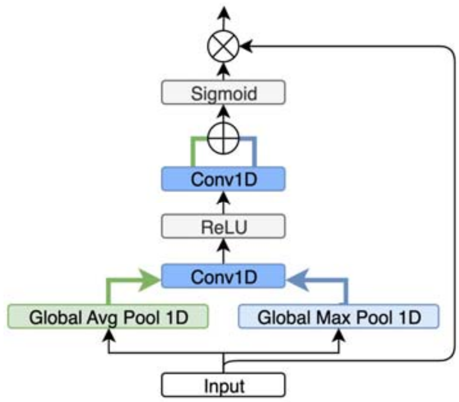

Figure 2 – Architecture of CA

Formally, the CA block used in this method is the following (1–3):

$$f_{avg} = Conv(RELU(Conv(AvgPool(f_{in}))), \quad (1)$$

$$f_{\max} = Conv(RELU(Conv(MaxPool(f_{in}))), \quad (2)$$

$$f_{out} = Sigmoid(f_{avg} + f_{\max}) \cdot f_{in} . \quad (3)$$

The MSFE (Fig. 3) module extracts preliminary features from the input signal at different scales, which makes the model more invariant to changes in the scale of fault occurrence. Inspired by [25], it consists of 3 branches, each enhanced by CA. The kernel sizes of each branch B_n in MSFE are: 301, 401, 501. Wide kernels help to reduce the influence of the noise that might be present in the input signal. The outputs of the branches are then concatenated and processed by the convolutional layer to effectively merge important information.

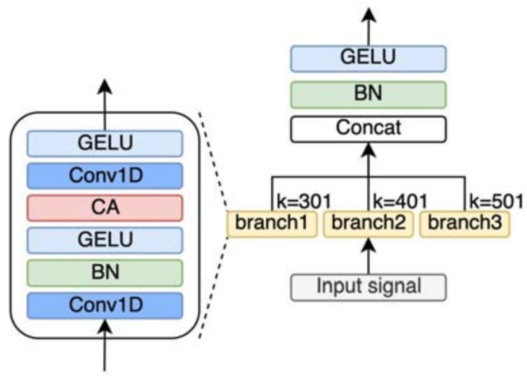

Figure 3 – Architecture of MSFE module

MSFE module is defined as follows (4):

$$f_{out} = GELU(BN(concat[B_1(f_{in}); B_2(f_{in}); B_3(f_{in})])) . \quad (4)$$

After the processing by MSFE, the signal is split into N_p patches before passed to the Encoder. First, each patch is encoded with convolutional embedding module to enrich them with more information. Then, positional encodings are added to the patches to preserve the information about the relative position of each patch. After that, the information is passed to the Encoder blocks.

The Encoder block (Fig. 4) follows the standard architecture [13] but with several changes. MHSA was replaced with CA, which treats each patch as a channel. It allows to significantly reduce the number of parameters while keeping the attention mechanism in the model. Position-wise FFN (Fig. 5) was made fully convolutional containing depth-wise convolution. These changes also reduce the complexity of the model while fusing the information from each patch.

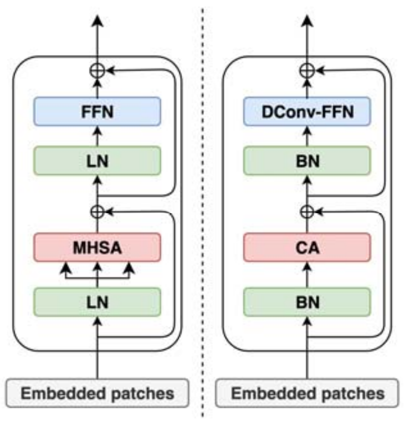

Figure 4 – A comparison of the vanilla Transformer block (left) and the proposed Transformer block (right)

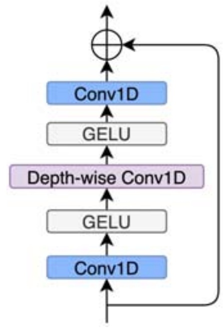

Figure 5 – Architecture of Position-wise FFN

Mathematically, the whole proposed Transformer block is defined as follows (5–6):

$$f^* = CA(BN(f_{in})) + f_{in} , \quad (5)$$

$$f_{out} = DconvFFN(BN(f^*)) + f^* . \quad (6)$$

Inspired by [31], the output of the Encoder is then passed through an average pooling layer and classification head to predict the type of the signal (healthy or faulty).

Table 1 describes the selected hyperparameters of the model that was used in the experiments.

Table 1 – The description of proposed model hyperparameters

Parameter	Value
Input size	1024
N_p	16
Patch embedding channels	64
Number of encoder blocks	6
FFN convolutional channels	64
Number of multi-scale branches	3
Branches' kernel sizes	(301, 401, 501)

4 EXPERIMENTS

To evaluate the performance of the proposed model, the aircraft fault dataset was used. This dataset contains 1158 samples of vibration signal, divided into two classes: 0 (healthy state) and 1 (fault). The amount of healthy and fault states is 865 and 293, respectively. Each sample is a vibration signal that contains 93752 datapoints. For training and testing, each sample was divided into chunks of length 1024 without overlapping, thus each sample provides 91 chunks. Each sample was scaled into the range of (–1, 1) before splitting into chunks.

To analyze the performance of the proposed module under limited data constraints, the number of samples in the training set was chosen to be much smaller than in the test set. More precisely, several training sets were created with 1, 2, 5, and 10 vibration signal samples of each class. The rest of the data was split between validation and testing sets. The information about created datasets is shown in Table 2. For example, the subset "Train 10" contains 20 randomly selected samples (10 samples for the positive class and 10 samples for the negative). Each sample was divided into chunks of 1024 datapoints, thus "Train 10" subset split contains 1820 chunks in total.

DOI 10.15588/1607-3274-2025-1-7

Table 2 – The distribution of the created data subsets

Dataset split	Samples	Classes (0/1)	Chunks
Train 1	2	1/1	182
Train 2	4	2/2	364
Train 5	10	5/5	910
Train 10	20	10/10	1820
Validation	113	89/24	10283
Test	1025	766/259	93275

To compare the proposed model with the existing methods, several models were chosen: MCNN-LSTM [35], WDCNN [17], MSCNN [37], MRACNN [38], MA1DCNN [39], ResNet18 (1D) [40].

Each model was trained for 30 epochs in 5 independent runs. After that, the results of each run were averaged. The batch size of the training phase is 128. AdamW was chosen as an optimization algorithm with a learning rate of 0.001. The loss function is binary cross-entropy.

To evaluate the accuracy of the proposed method under a noisy environment, the Gaussian white noise was added to the raw signal to get the signal with different SNR (7) values. In these experiments, noise with SNR from –9dB to 9dB with the step of 3dB was added.

$$SNR = 10\log_{10}\frac{P_s}{P_n}. \quad (7)$$

For comparing results, AP was chosen as a metric. AP is a metric commonly used in binary classification, especially when dataset is highly imbalanced. It is essentially the area under the precision-recall curve, providing a single score that summarize the precision-recall trade-off across various thresholds. The range of AP metric values is (0–1), where 1 means errorless classification.

5 RESULTS

Table 3 shows the results on the test dataset of different methods. Each model is trained on datasets with a different number of vibration samples to evaluate the performance in limited data conditions. Also, the number of parameters of each model is included. The results represent the value of AP metric. The results show that the proposed method (LMCT) trained on a subset of 20 samples (“Train 10”) outperforms the chosen existing methods (AP 0.9941) while having a relatively small number of parameters (0.097M).

Table 3 – The prediction results of selected methods trained on different dataset sizes

Method	N samples per class				N params (M)
	1	2	5	10	
MCNN-LSTM	0.3154	0.4016	0.4342	0.5421	0.093
WDCNN	0.3455	0.6399	0.8322	0.9216	0.041
MSCNN	0.3337	0.7790	0.7998	0.9511	13.78
MRACNN	0.3372	0.5732	0.9182	0.9699	0.599
MA1DCNN	0.3174	0.8043	0.8626	0.9094	0.323
Resnet18 (1D)	0.3346	0.6631	0.8982	0.9556	3.84
LMCT (proposed)	0.3133	0.7314	0.9105	0.9941	0.097

Note that these results represent the performance of the models on the dataset that is collected from samples’ chunks. To analyze the whole vibration sample, the average prediction of each chunk can be used.

Table 4 shows the ablation study results to evaluate the performance and the importance of proposed changes, especially the MSFE module and new Convolutional Transformer Encoder block. The baseline model, the 1D Transformer, follows the standard Transformer architecture. The hyperparameters in the baseline were chosen to make it as close to the proposed method as possible.

Table 4 – The ablation study of the effect of proposed improvements

Method	N samples per class			
	1	2	5	10
1D Transformer	0.3122	0.2976	0.2963	0.3919
1D Transformer + MSFE	0.3122	0.3904	0.4869	0.6063
LMCT w/o MSFE	0.3137	0.6335	0.7523	0.8056
LMCT	0.3133	0.7314	0.9105	0.9941

Figure 6 shows the results (AP) of the selected models under noisy environment conditions (SNR from –9 to 9). For that, models that were trained on 20 samples (subset “Train 10”) were chosen.

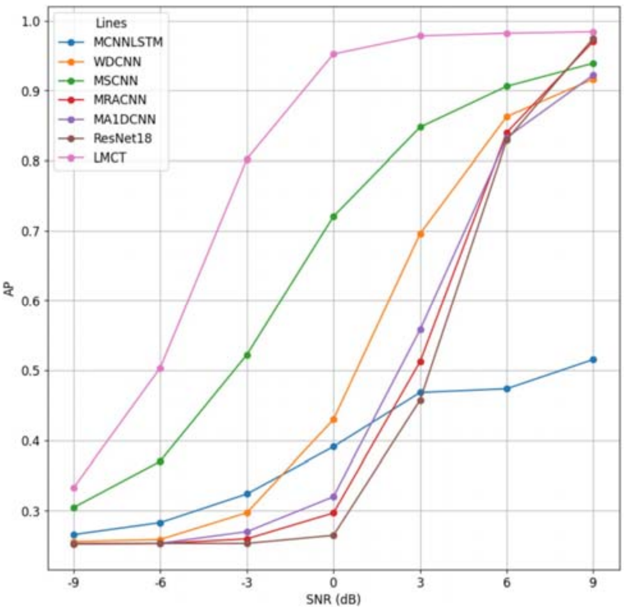

Figure 6 – The results under different SNRs of Gaussian noise

The visualization of the learnt features (featurs from the last layer of the network) using t-SNE method is show in the Fig. 7.

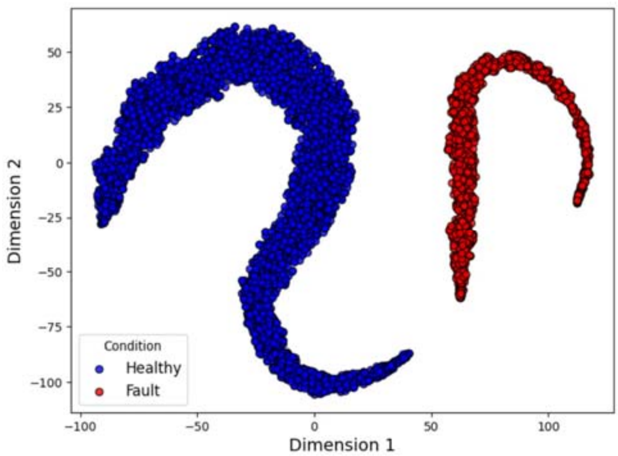

Figure 7 – Distribution of features visualized with t-SNE

DOI 10.15588/1607-3274-2025-1-7

6 DISCUSSION

The results from Table 3 show that the proposed model achieves almost ideal accuracy (AP 0.9941). Moreover, the number of parameters of the proposed model is relatively small (0.091M) compared to the existing approaches, which means that a small model is able to achieve good results in vibration-based FD tasks. This can be explained by the fact that the proposed methods employ a combination of convolution and attention mechanisms, which allows the model to examine local features as well as long-term dependencies.

The results of the ablation study in Table 4 suggest the importance of the proposed modules. The MSFE module significantly improves the performance, which means that it is able to extract valuable deep features from the input signal. The Transformer backbone is able to process these features more accurately than the raw input signal. Also, the vanilla Transformer model is not able to accurately classify aircraft sensor vibration signals. It can be explained by the fact that each signal can contain local patterns that can influence the overall prediction results. To extract these patterns, the strength of convolutional inductive locality bias is needed.

Figure 6 shows that the proposed model can be used in noisy environments effectively. The proposed model achieves AP of 0.9783, 0.9821, and 0.9842 at SNR values of 3dB, 6dB, and 9dB, respectively. When SNR is higher than 0, it means that the energy of the original signal is higher than the energy of the added noise and vice versa. In cases when the energy of the noise is higher than the energy of the original signal, classification is difficult to perform. The results, obtained under different SNR values, indicate that wide kernels in MSFE help the model to reduce the impact of the noise in the signal.

Figure 7, which shows the 2D representation of extracted features, suggests that the model is able to learn the feature distribution by extracting patterns and dependencies between similar entries. The fault features and healthy features are clustered, which means that model can separate between both clusters.

CONCLUSIONS

In this paper, the problem of fault diagnosis of aircraft rotating machinery is being solved by applying deep learning method.

The scientific novelty. The conducted experiments show that proposed method can achieve high accuracy having small number of parameters. The proposed changes made to Transformer architecture significantly increase model's performance while reducing it's size. Moreover, the results show that FD methods can be effectively trained with limited size of the data which is important in the field of aircraft FD where fault data samples are extremely rare.

The practical significance. The obtained model allows diagnosing faults in aircraft with high accuracy and can be applied to other rotating machinery FD tasks that use vibration-based analysis.

The prospects for further research lie in developing more efficient model and testing with other limited data training techniques like zero-shot or few-shot learning.

ACKNOWLEDGEMENTS

This work was performed within the scope of the All-Ukrainian competition-hackathon of scientific works of young scientists in the field of intellectual information technologies, 2024. The data was provided by Bechhoefer Eric, PhD, GPMS Inc. (Cornwall, USA).

REFERENCES

1. Baccarini L. M. R., Silva V. V. R., de Menezes B. R. et al. SVM practical industrial application for mechanical faults diagnostic, *Expert Systems with Applications,* 2011, Vol. 38, № 6, pp. 6980–6984. DOI: 10.1016/j.eswa.2010.12.017
2. Lu Q., Shen X., Wang X. et al. Fault diagnosis of rolling bearing based on improved VDM and KNN, *Mathematical Problems in Engineering*, 2021, Vol. 2021, Issue 1, №2530315. DOI: 10.1155/2021/2530315
3. Sun W., Chen J., Li J.Decision tree and PCA-based fault diagnosis of rotating machinery, M*echanical Systems and Signal Processing*, 2007, Vol. 21, Issue 3, pp. 1300–1317. DOI: 10.1016/j.ymssp.2006.06.010
4. Zhao R., Yan R., Chen Z. et al. Deep learning and its applications to machine health monitoring, *Mechanical Systems and Signal Processing*, 2019, Vol. 115, pp. 213–237. DOI: 10.1016/j.ymssp.2018.05.050
5. Samanta B., Al-Balushi K., and Al-Araimi S. Artificial neural networks and genetic algorithm for bearing fault detection, *Soft Computing*, 2006, Vol. 10, pp. 264–271. DOI: 10.1007/s00500-005-0481-0
6. Wang X., Mao D., Li X. Bearing fault diagnosis based on vibro-acoustic data fusion and 1D-CNN network, *Measurement,* 2021, Vol. 173, №108518.DOI: 10.1016/j.measurement.2020.108518
7. Pinedo-Sanchez L. A., Mercado-Ravell D. A., Carballo-Monsivais C. A. Vibration analysis in bearings for failure prevention using CNN, *Journal of the Brazilian Society of Mechanical Sciences and Engineering,* 2020, Vol. 42, № 12, № 628. DOI: 10.1007/s40430-020-02711-w
8. Yu Z., Zhang C., Liu J. et al. SKND-TSACNN: A novel time-scale adaptive CNN framework for fault diagnosis of rotating machinery, *Knowledge-Based Systems,* 2023, Vol. 275, P. 110682. DOI: 10.1016/j.knosys.2023.110682;
9. Huang Y., Chen C. H., Huang C. J. Motor fault detection and feature extraction using RNN-based variational autoencoder, *IEEE Access,* 2019, Vol. 7, pp. 139086–139096. DOI: 10.1109/ACCESS.2019.2940769
10. Liu H., Zhou J., Zheng Y. et al. Fault diagnosis of rolling bearings with recurrent neural network-based autoencoders, *ISA Transactions,* 2018, Vol. 77, pp. 167–178. DOI: 10.1016/j.isatra.2018.04.005
11. Saghi T., Bustan D., Aphale S. S. Bearing fault diagnosis based on multi-scale CNN and bidirectional GRU, *Vibration*, 2023, Vol. 6, № 1, pp. 11–28. DOI: 10.3390/vibration6010002
12. Qiao M., Yan S., Tang X. et al. Deep Convolutional and LSTM Recurrent Neural Networks for Rolling Bearing Fault Diagnosis Under Strong Noises and Variable Loads, *IEEE Access,* 2020, Vol. 8, pp. 66257–66269. DOI: 10.1109/ACCESS.2020.2985617
13. Vaswani A., Shazeer N., Parmar N. et al. Attention is All you Need [Electronic resource], *Semantic Schoolar,* 2017,

DOI 10.15588/1607-3274-2025-1-7

Access mode: https://www.semanticscholar.org/paper/Attention-is-Allyou-Need-Vaswani-Shazeer/204e3073870fae3d05bcbc2f6a8e263d9b72e776. DOI: 10.48550/arXiv.1706.03762
14. Xu J., L. Zhou, W. Zhao et al. Zero-shot learning for compound fault diagnosis of bearings, *Expert Systems with Applications,* 2022, Vol. 190, №116197. DOI: 10.1016/j.eswa.2021.116197
15. Zhang A., Li S., Cui Y. et al. Limited Data Rolling Bearing Fault Diagnosis With Few-Shot Learning, *IEEE Access*, 2019, Vol. 7, pp. 110895–110904. DOI: 10.1109/ACCESS.2019.2934233
16. Chen C. C., Liu Z., Yang G. et al. An improved fault diagnosis using 1D-convolutional neural network model, *Electronics,* 2021, Vol. 10, Issue 1, pp. 1–19. DOI: 10.3390/electronics10010059
17. Zhang W., Peng G., Li C. et al. A New Deep Learning Model for Fault Diagnosis with Good Anti-Noise and Domain Adaptation Ability on Raw Vibration Signals, *Sensors*, 2017, Vol. 17, Issue 2, №425. DOI: 10.3390/s17020425
18. Zhao M., Zhong S., Fu X. et al. Deep Residual Shrinkage Networks for Fault Diagnosis, *IEEE Transactions on Industrial Informatics*, 2020, Vol. 16, Issue 7, pp. 4681–4690. DOI: 10.1109/TII.2019.2943898
19. Zhang W., Li X., Ding Q. Deep residual learning-based fault diagnosis method for rotating machinery, *ISA Transactions,* 2019, Vol. 95, pp. 295–305. DOI: 10.1016/j.isatra.2018.12.025
20. Li X., Wan S., Liu S. et al. Bearing fault diagnosis method based on attention mechanism and multilayer fusion network, *ISA Transactions*, 2021, Vol. 128, Part B, pp. 550–564. DOI: 10.1016/j.isatra.2021.11.020
21. Verstraete D., Ferrada A., Droguett E. et al. Deep Learning Enabled Fault Diagnosis Using Time-Frequency Image Analysis of Rolling Element Bearings, *Shock and Vibration,* 2017, Vol. 2017, Issue 1, №5067651. DOI: 10.1155/2017/5067651
22. Pham M. T., Kim J.-M., and Kim C. H. Accurate Bearing Fault Diagnosis under Variable Shaft Speed using Convolutional Neural Networks and Vibration Spectrogram, *Applied Sciences,* 2020, Vol. 10, Issue 18, № 6385. DOI: 10.3390/app10186385;
23. Dosovitskiy A., Beyer L., Kolesnikov A. et al. An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale [Electronic resource], *ICLR,* 2021. Access mode: https://arxiv.org/abs/2010.11929. DOI: 10.48550/arXiv.2010.11929
24. Zhou H., Zhang S., Peng J. et al. Informer: Beyond Efficient Transformer for Long Sequence Time-Series Forecasting, *Proceedings of the AAAI conference on artificial intelligence,* 2021, Vol. 35, Issue 12, pp. 11106–11115. DOI: 10.48550/arXiv.2012.07436
25. Weng C., Lu B., Yao J. A One-Dimensional Vision Transformer with Multiscale Convolution Fusion for Bearing Fault Diagnosis, *Global Reliability and Prognostics and Health Management*. N., IEEE, 2021, pp. 1–6. DOI: 10.1109/PHM-Nanjing52125.2021.9612919
26. Jin Y., Hou L., Chen Y. A Time Series Transformer based method for the rotating machinery fault diagnosis, *Neurocomputing*, 2022, Vol. 494, pp. 379–395. DOI: 10.1016/j.neucom.2022.04.111
27. Zim A. H., Ashraf A., Iqbal A. et al. A Vision Transformer-Based Approach to Bearing Fault Classification via Vibration Signals, *Asia-Pacific Signal and Information Processing Association Annual Summit and Conference*. C., IEEE, 2022, pp. 1321–1326. DOI: 10.48550/arXiv.2208.07070
28. Liu W., Zhang Z., Zhang J. et al. A Novel Fault Diagnosis Method of Rolling Bearings Combining Convolutional Neural Network and Transformer, *Electronics,* 2023,Vol. 12, Issue 8, №1838. DOI: 10.3390/electronics12081838
29. Tang X., Xu Z., Wang Z. A Novel Fault Diagnosis Method of Rolling Bearing Based on Integrated Vision Transformer Model, *Sensors,* 2022, Vol. 22, Issue 10, №3878. DOI: 10.3390/s22103878
30. Xu Y., Zhang Q., Zhang J. et al. ViTAE: vision transformer advanced by exploring intrinsic inductive bias, *NIPS'21 : 35th International Conference on Neural Information Processing Systems 6–14 December 2021 : proceedings*. Red Hook, NY, Curran Associates Inc., 2021, Vol. 34, pp. 28522–28535. DOI: 10.48550/arXiv.2106.03348
31. Sun W., Yan R., Jin R. et al. LiteFormer: A Lightweight and Efficient Transformer for Rotating Machine Fault Diagnosis, *IEEE Transactions on Reliability,* 2024, Vol. 73, Issue 2, pp. 1258–1269. DOI: 10.1109/TR.2023.3322860
32. Fang X., Deng X., Chen J. et al. Convolution Transformer Based Fault Diagnosis Method For Aircraft Engine Bearings, *43rd Chinese Control Conference, Kunming, 28–31 July 2024 : proceedings*. IEEE, 2024, pp. 5038–5042. DOI: 10.23919/CCC63176.2024.10662769
33. Pei X., Zheng X., Wu J. Rotating Machinery Fault Diagnosis Through a Transformer Convolution Network Subjected to Transfer Learning, *IEEE Transactions on Instrumentation and Measurement*, 2021, Vol. 70, №2515611. DOI: 10.1109/TIM.2021.3119137
34. Jin Y., Qin C., Zhang Z. et al. A multi-scale convolutional neural network for bearing compound fault diagnosis under various noise conditions, *Science China Technological Sciences,* 2022, Vol. 65, Issue 11, pp. 2551–2563. DOI: 10.1007/s11431-022-2109-4
35. Chen X., Zhang B., Gao D. Bearing fault diagnosis base on multi-scale CNN and LSTM model, *Journal of Intelligent Manufacturing*, 2021, Vol. 32, Issue 4, pp. 971–987. DOI: 10.1007/s10845-020-01600-2
36. Hu J., Shen L., Sun G. Squeeze-and-Excitation Networks, *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, 18–23 June 2018 : preceedings*. IEEE, 2018, pp. 7132–7141. DOI: 10.1109/CVPR.2018.00745
37. Jiang G., He H., Yan J. et al. Multiscale Convolutional Neural Networks for Fault Diagnosis of Wind Turbine Gearbox, *IEEE Transactions on Industrial Electronics*, 2019, Vol. 66, Issue 4, pp. 3196–3207. DOI: 10.1109/TIE.2018.2844805;
38. Jia L., Chow T. W. S., Wang Y. et al. Multiscale Residual Attention Convolutional Neural Network for Bearing Fault Diagnosis, *IEEE Transactions on Instrumentation and Measurement*, 2022, Vol. 71, №3519413. DOI: 10.1109/TIM.2022.3196742
39. Wang H., Liu Z., Peng D. et al. Understanding and Learning Discriminant Features based on Multiattention 1DCNN for Wheelset Bearing Fault Diagnosis, *IEEE Transactions on Industrial Informatics*, 2020, Vol. 16, Issue 9, pp. 5735–5745. DOI: 10.1109/TII.2019.2955540
40. He K., Zhang X., Ren S. et al. Deep Residual Learning for Image Recognition, *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, 27–30 June 2016.* IEEE, 2016, pp. 770–778, DOI: 10.1109/CVPR.2016.90

Received 09.12.2024.
Accepted 03.02.2025.

DOI 10.15588/1607-3274-2025-1-7

УДК 004.93

ЛЕГКОВІСНИЙ БАГАТОМАСШТАБНИЙ ЗГОРТКОВИЙ ТРАНСФОРМЕР ДЛЯ ДІАГНОСТИКИ НЕСПРАВНОСТЕЙ ЛІТАЛЬНИХ АПАРАТІВ ЗА ДОПОМОГОЮ ВІБРАЦІЙНОГО АНАЛІЗУ

Діденко Андрій Є. – аспірант кафедри програмних засобів національного університету «Запорізька Політехніка», Запоріжжя, Україна.

Діденко Артем Є. – аспірант кафедри програмних засобів національного університету «Запорізька Політехніка», Запоріжжя, Україна.

Субботін С. О. – д-р техн. наук, професор, завідувач кафедри програмних засобів національного університету «Запорізька Політехніка», Запоріжжя, Україна.

АНОТАЦІЯ

Актуальність. Діагностика несправностей обертових механізмів, особливо в авіації, відіграє важливу роль в системах моніторингу стану. Своєчасне і точне виявлення несправностей може значно знизити вартість ремонту і збільшити термін служби механізму. Для ефективного виявлення несправностей використовуються інтелектуальні методи, які базуються на традиційних методах машинного та глибинного навчання. Об'єктом дослідження є процес виявлення несправностей в авіаційних апаратах на основі аналізу вібрацій.

Метою роботи є розробка методу глибинног навчання для діагностики несправностей обертових машин з високою точністю.

Метод. Запропонований метод використовує архітектуру трансформера. Першим етапом обробки сигналу вібрації є багатомасштабне вилучення ознак. Цей етап дозволяє моделі розглядати вхідні сигнали в різних масштабах і зменшити впливи шуму. Другий етап – згорткова нейронна мережа з трансормером. Згортка була додана до трансформеру, щоб поєднати локальність і вилучення ознак далеких залежностей. Механізм самоуваги трансормеру було змінено на механізм канальної уваги, що зменшує кількість параметрів, але зберігає силу уваги. Щоб підсилити цю ідею, аналогічні зміни були зроблені в позиційній мережі прямого поширення.

Результати. Запропонований метод протестовано на наборі даних з вібраціями авіаційного апарату. Для тестування було обрано дві умови: обмеженість обсягу даних та зашумлене середовище. Обмеженість обсягу даних імітується шляхом використання невелої кількості вибірок до навчального набору даних (максимум 10 на клас). Умова зашумленого середовища імітується шляхом додавання гауссівського шуму до вихідного сигналу. Згідно з отриманими результатами, запропонований метод досягає високої середньої точності при невеликій кількості параметрів. Експерименти також показують важливість запропонованих модулів і змін, підтверджуючи припущення про процес вилучення ознак.

Висновки. Результати проведених експериментів показують, що запропонована модель може виявляти несправності з майже ідеальною точністю, навіть при невеликій кількості параметрів. Запропонована легковісна модель є стійкою в умовах обмеженого обсягу даних та зашумленого середовища. Перспективами подальших досліджень є розробка швидких і точних нейронних мереж для діагностики несправностей та розробка методів навчання на обмежених обсягах даних.

КЛЮЧОВІ СЛОВА: аналіз несправностей, глибинне навчання, обертові механізми, обробка сигналів, трансформер, нейронні мережі.

ЛІТЕРАТУРА

1. SVM practical industrial application for mechanical faults diagnostic / [L. M. R. Baccarini, V. V. R. Silva, B. R. de Menezes et al.] // Expert Systems with Applications. – 2011. – Vol. 38, № 6. – P. 6980–6984. DOI: 10.1016/j.eswa.2010.12.017
2. Fault diagnosis of rolling bearing based on improved VDM and KNN / [Q. Lu, X. Shen, X. Wang et al.] // Mathematical Problems in Engineering. – 2021. – Vol. 2021, Issue 1. – №2530315. DOI: 10.1155/2021/2530315
3. Sun W. Decision tree and PCA-based fault diagnosis of rotating machinery / W. Sun, J. Chen, J. Li // Mechanical Systems and Signal Processing. – 2007. – Vol. 21, Issue 3. – P. 1300–1317. DOI: 10.1016/j.ymssp.2006.06.010
4. Deep learning and its applications to machine health monitoring / [R. Zhao, R. Yan, Z. Chen et al.] // Mechanical Systems and Signal Processing. – 2019. – Vol. 115. – P. 213–237. DOI: 10.1016/j.ymssp.2018.05.050
5. Samanta B. Artificial neural networks and genetic algorithm for bearing fault detection / B. Samanta, K. Al-Balushi, and S. Al-Araimi // Soft Computing. – 2006. – Vol. 10. – P. 264–271. DOI: 10.1007/s00500-005-0481-0
6. Wang X. Bearing fault diagnosis based on vibro-acoustic data fusion and 1D-CNN network / X. Wang, D. Mao, X. Li // Measurement. – 2021. – Vol. 173. – №108518.DOI: 10.1016/j.measurement.2020.108518
7. Pinedo-Sanchez L. A. Vibration analysis in bearings for failure prevention using CNN / L. A. Pinedo-Sanchez, D. A. Mercado-Ravell, C. A. Carballo-Monsivais // Journal of the Brazilian Society of Mechanical Sciences and Engineering. – 2020. – Vol. 42, № 12. – №628. DOI: 10.1007/s40430-020-02711-w
8. SKND-TSACNN: A novel time-scale adaptive CNN framework for fault diagnosis of rotating machinery / [Z. Yu, C. Zhang, J. Liu et al.] // Knowledge-Based Systems. – 2023. – Vol. 275. – P. 110682. DOI: 10.1016/j.knosys.2023.110682
9. Huang Y. Motor fault detection and feature extraction using RNN-based variational autoencoder / Y. Huang, C. H. Chen, C. J. Huang // IEEE Access. – 2019. – Vol. 7. – P. 139086–139096. DOI: 10.1109/ACCESS.2019.2940769
10. Fault diagnosis of rolling bearings with recurrent neural network-based autoencoders / [H. Liu, J. Zhou, Y. Zheng et al.] // ISA Transactions. – 2018. – Vol. 77. – P. 167–178. DOI: 10.1016/j.isatra.2018.04.005
11. Saghi T. Bearing fault diagnosis based on multi-scale CNN and bidirectional GRU / T. Saghi, D. Bustan, S. S. Aphale // Vibration. – 2023. – Vol. 6, № 1. – P. 11–28. DOI: 10.3390/vibration6010002
12. Deep Convolutional and LSTM Recurrent Neural Networks for Rolling Bearing Fault Diagnosis Under Strong Noises and Variable Loads / [M. Qiao, S. Yan, X. Tang et al.] // IEEE Access. – 2020. – Vol. 8. – P. 66257–66269. DOI: 10.1109/ACCESS.2020.2985617
13. Attention is All you Need [Electronic resource] / [A. Vaswani, N. Shazeer, N. Parmar et al.] // Semantic Schoolar. –

DOI 10.15588/1607-3274-2025-1-7

2017. – Access mode: https://www.semanticscholar.org/paper/Attention-is-Allyou-Need-Vaswani-Shazeer/204e3073870fae3d05bcbc2f6a8e263d9b72e776. DOI: 10.48550/arXiv.1706.03762
14. Zero-shot learning for compound fault diagnosis of bearings / [J. Xu, L. Zhou, W. Zhao et al.] // Expert Systems with Applications. – 2022. – Vol. 190. – №116197. DOI: 10.1016/j.eswa.2021.116197
15. Limited Data Rolling Bearing Fault Diagnosis With Few-Shot Learning / [A. Zhang, S. Li, Y. Cui et al.] // IEEE Access. – 2019. – Vol. 7. – P. 110895–110904. DOI: 10.1109/ACCESS.2019.2934233
16. An improved fault diagnosis using 1D-convolutional neural network model / [C. C. Chen, Z. Liu, G. Yang et al.] // Electronics. – 2021. – Vol. 10, Issue 1. – P. 1–19. DOI: 10.3390/electronics10010059
17. A New Deep Learning Model for Fault Diagnosis with Good Anti-Noise and Domain Adaptation Ability on Raw Vibration Signals / [W. Zhang, G. Peng, C. Li et al.] // Sensors. – 2017. – Vol. 17, Issue 2. – №425. DOI: 10.3390/s17020425
18. Deep Residual Shrinkage Networks for Fault Diagnosis / [M. Zhao, S. Zhong, X. Fu et al.] // IEEE Transactions on Industrial Informatics. – 2020. – Vol. 16, Issue 7. – P. 4681–4690. DOI: 10.1109/TII.2019.2943898
19. Zhang W. Deep residual learning-based fault diagnosis method for rotating machinery / W. Zhang, X. Li, Q. Ding // ISA Transactions. – 2019. – Vol. 95. – P. 295–305. DOI: 10.1016/j.isatra.2018.12.025
20. Bearing fault diagnosis method based on attention mechanism and multilayer fusion network / [X. Li, S. Wan, S. Liu et al.] // ISA Transactions. – 2021. – Vol. 128, Part B. – P. 550–564. DOI: 10.1016/j.isatra.2021.11.020
21. Deep Learning Enabled Fault Diagnosis Using Time-Frequency Image Analysis of Rolling Element Bearings / [D. Verstraete, A. Ferrada, E. Droguett et al.] // Shock and Vibration. – 2017. – Vol. 2017, Issue 1. – №5067651. DOI: 10.1155/2017/5067651
22. Pham M. T. Accurate Bearing Fault Diagnosis under Variable Shaft Speed using Convolutional Neural Networks and Vibration Spectrogram / M. T. Pham, J.-M. Kim, and C. H. Kim // Applied Sciences. – 2020. – Vol. 10, Issue 18. – №6385. DOI: 10.3390/app10186385
23. An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale [Electronic resource] / [A. Dosovitskiy, L. Beyer, A. Kolesnikov et al.] // ICLR. – 2021. – Access mode: https://arxiv.org/abs/2010.11929. DOI: 10.48550/arXiv.2010.11929
24. Informer: Beyond Efficient Transformer for Long Sequence Time-Series Forecasting / [H. Zhou, S. Zhang, J. Peng et al.] // Proceedings of the AAAI conference on artificial intelligence. – 2021. – Vol. 35, Issue 12. – P. 11106–11115. DOI: 10.48550/arXiv.2012.07436
25. Weng. C. A One-Dimensional Vision Transformer with Multiscale Convolution Fusion for Bearing Fault Diagnosis / C. Weng, B. Lu, J. Yao // Global Reliability and Prognostics and Health Management. – N. : IEEE, 2021. – P. 1–6. DOI: 10.1109/PHM-Nanjing52125.2021.9612919
26. Jin Y. A Time Series Transformer based method for the rotating machinery fault diagnosis / Y. Jin, L. Hou, Y. Chen // Neurocomputing. – 2022. – Vol. 494. – P. 379–395. DOI: 10.1016/j.neucom.2022.04.111
27. A Vision Transformer-Based Approach to Bearing Fault Classification via Vibration Signals / [A. H. Zim, A. Ashraf, A. Iqbal et al.] // Asia-Pacific Signal and Information Processing Association Annual Summit and Conference. – C.:IEEE, 2022. – P. 1321–1326. DOI: 10.48550/arXiv.2208.07070
28. A Novel Fault Diagnosis Method of Rolling Bearings Combining Convolutional Neural Network and Transformer / [W. Liu, Z. Zhang, J. Zhang et al.] // Electronics. – 2023. – Vol. 12, Issue 8. – №1838. DOI: 10.3390/electronics12081838
29. Tang X. A Novel Fault Diagnosis Method of Rolling Bearing Based on Integrated Vision Transformer Model / X. Tang, Z. Xu, Z. Wang // Sensors. – 2022. – Vol. 22, Issue 10. – №3878. DOI: 10.3390/s22103878
30. ViTAE: vision transformer advanced by exploring intrinsic inductive bias / [Y. Xu, Q. Zhang, J. Zhang et al.] // NIPS'21 : 35th International Conference on Neural Information Processing Systems 6–14 December 2021 : proceedings. – Red Hook, NY : Curran Associates Inc., 2021. – Vol. 34. – P. 28522–28535. DOI: 10.48550/arXiv.2106.03348
31. LiteFormer: A Lightweight and Efficient Transformer for Rotating Machine Fault Diagnosis / [W. Sun, R. Yan, R. Jin et al.] // IEEE Transactions on Reliability. – 2024. – Vol. 73, Issue 2, P. 1258–1269. DOI: 10.1109/TR.2023.3322860
32. Convolution Transformer Based Fault Diagnosis Method For Aircraft Engine Bearings / [X. Fang, X. Deng, J. Chen et al.] // 43rd Chinese Control Conference, Kunming, 28–31 July 2024 : proceedings. – IEEE, 2024. – P. 5038–5042. DOI: 10.23919/CCC63176.2024.10662769
33. Pei X. Rotating Machinery Fault Diagnosis Through a Transformer Convolution Network Subjected to Transfer Learning / X. Pei, X. Zheng, J. Wu // IEEE Transactions on Instrumentation and Measurement. – 2021. – Vol. 70. – №2515611. DOI: 10.1109/TIM.2021.3119137
34. A multi-scale convolutional neural network for bearing compound fault diagnosis under various noise conditions / [Y. Jin, C. Qin, Z. Zhang et al.] // Science China Technological Sciences. – 2022. – Vol. 65, Issue 11. – P. 2551–2563. DOI: 10.1007/s11431-022-2109-4
35. Chen X. Bearing fault diagnosis base on multi-scale CNN and LSTM model / X. Chen, B. Zhang, D. Gao // Journal of Intelligent Manufacturing. – 2021. – Vol. 32, Issue 4. – P. 971–987. DOI: 10.1007/s10845-020-01600-2
36. Hu J. Squeeze-and-Excitation Networks / J. Hu, L. Shen, G. Sun // 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, 18–23 June 2018 : preceedings. – IEEE, 2018. – P. 7132–7141. DOI: 10.1109/CVPR.2018.00745
37. Multiscale Convolutional Neural Networks for Fault Diagnosis of Wind Turbine Gearbox / [G. Jiang, H. He, J. Yan et al.] // IEEE Transactions on Industrial Electronics– 2019. – Vol. 66, Issue 4. – P. 3196–3207. DOI: 10.1109/TIE.2018.2844805
38. Multiscale Residual Attention Convolutional Neural Network for Bearing Fault Diagnosis / [L. Jia, T. W. S. Chow, Y. Wang et al.] // IEEE Transactions on Instrumentation and Measurement. – 2022. – Vol. 71. – №3519413. DOI: 10.1109/TIM.2022.3196742
39. Understanding and Learning Discriminant Features based on Multiattention 1DCNN for Wheelset Bearing Fault Diagnosis / [H. Wang, Z. Liu, D. Peng et al.] // IEEE Transactions on Industrial Informatics. – 2020. – Vol. 16, Issue 9. – P. 5735–5745. DOI: 10.1109/TII.2019.2955540
40. Deep Residual Learning for Image Recognition / [K. He, X. Zhang, S. Ren et al.] // 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, 27–30 June 2016. – IEEE, 2016. – P. 770–778, DOI: 10.1109/CVPR.2016.90

DOI 10.15588/1607-3274-2025-1-7