

Deep Stacked Auto-Encoder Network Based Tool Wear Monitoring in the Face Milling Process

VanThien Nguyen – VietHung Nguyen* – VanTrinh Pham
Hanoi University of Industry, VietNam

Tool wear identification plays an important role in improving product quality and productivity in the manufacturing industry. The actual tool wear status with input cutting parameters may cause different levels of spindle vibration during the machining process. This research proposes an architecture comprising a deep learning network (DLN) to identify the actual wear state of machining tool. Firstly, data on spindle vibration signals are obtained from an acceleration sensor. The data are then pre-processed using the fast Fourier transform (FFT) method to reveal the relevant outstanding features in the frequency domain. Finally, the DLN is constructed based on stacked auto-encoders (SAE) and softmax, which is trained with the input data on the vibration features of the respective tool wear state. This DLN architecture is then used to identify the actual wear statuses of machining tool. The experimental results from the collected data show that the proposed DLN architecture is capable of identifying actual tool wear with high accuracy.

Keywords: Face milling; Tool wear; Stacked auto-encoder (SAE); Deep learning network (DLN); Cast iron

Highlights

- An expert technique for tool wear monitoring based on an experimental dataset is explored.
- The feature values with respect to the wear status of cutting tools are extracted and analyzed.
- The effects of a proposed deep learning network architecture for identifying the different tool wear statuses are considered.
- A patterns prediction method is compared and developed.

0 INTRODUCTION

In the machining process, cutting tool play an important role and can affect the stability of machine components and systems. Cutting tool status is closely related to the vibration, cutting heat and cutting force of a machining system, which normally includes the machine tool, cutting tool, workpiece, and clamping device. The working status of a cutting tool significantly influences machining efficiency and thus the accuracy of a product [1] to [3]. Therefore, monitoring cutting tool status is important. Normally, tool wear can cause major failure of the cutting tool, which may result in wastage of a product, slow operation, and a lack of productivity. Identifying the wear status of cutting tool is especially important in evaluating product accuracy [4] and [5].

The features of the original vibration signal can provide useful information for assessing the wear condition of cutting tool and can be extracted by principal component analysis (PCA) [6] and linear discriminate analysis (LDA) [7] and [6]. Dimensional feature data, in particular, can cause a computational burden, affect the efficiency of the classification phase as it is time-consuming, and reduce the diagnosis accuracy. These methods are effective with linear data but cannot be effectively applied for complex nonlinear and nonstationary vibration data.

Machine learning techniques have been researched and applied in many fields of science and technology, such as computer vision, automatic diagnostics systems and pattern recognition. These techniques are often based on artificial intelligence methods such as k- nearest neighbour (K-NN) [8], support vector machines (SVM) [9] and [10], and the artificial neural network (ANN) [11] to [13]. These methods have been effectively applied to identify tool wear status. However, they have some weaknesses, including ineffective feature extraction. In particular, the objectivity of unsupervised feature learning has been ignored, and automation is not employed in these methods [14] and [15]. Therefore, using these techniques to identify tool wear in the machining process has been unsatisfactory.

Recently, deep learning network (DLN) architecture has been widely used in research [16] and some engineering science applications such as medical informatics [17], pattern recognition [18] and [19], and time-series prediction [20] to [22]. DLNs are hierarchically constructed with many hidden layers with the aim of effectiveness in the output layer [23]. In this study, the authors propose a DLN architecture based on the SAE and softmax classifier, which are closely stacked together to implement tool wear diagnosis in the end-milling process. In this architecture, each auto-encoder (AE) implements vibration data reconstruction to generate higher-

level features with an unsupervised algorithm that is optimized to minimize errors between the input and output data. The relevant vibration features are precisely extracted at the hidden layer of the last AE, which can significantly improve the learning effect of the classification phase. In addition, softmax is derived from the multinomial logistic model, which is based on the supervised learning algorithm and is suitable for multiclass classification [24]. Finally, the parameters of the proposed DLN will be fine-tuned with the supervised condition in its complete architecture, with the goal of effectiveness of the classification accuracy result.

The rest of the paper is structured as follows: Section 1 presents the materials. Section 2 proposes a DLN architecture, which is then used to identify tool wear status of the end-milling process. Section 3 presents experimental results and discusses the diagnosis results, which are analyzed and compared. Section 4 is the conclusion. Acknowledgments and a list of references then follow.

1 MATERIALS

This section expresses the single AE architecture and softmax classifier model that are the basis for constructing the DLN for the diagnosis technique. The AE is used to exploit the features of the original vibration signal related to tool wear status. These features are then used to train softmax classifier and construct the proposed DLN.

1.1 Auto-Encoder Network Architecture

In [23], AE is a special type of neural network architecture that acts as an unsupervised algorithm. The AE architecture consists of three layers instead of the input layer, the hidden layer and the output layer, which are organized into two phases -encoding and decoding- as shown in Fig. 1. The input layer $x = \{x_1, x_2, \dots, x_n\}$, hidden layer $f = \{f_1, f_2, \dots, f_m\}$, $m \ll n$, and output layer $\tilde{x} = \{\tilde{x}_1, \tilde{x}_2, \dots, \tilde{x}_n\}$ are seamlessly connected. This AE is implemented to reconstruct data of the input layer.

The encoder phase has encoded the characteristics of the high-dimensional input data x into the low-dimensional f data in the hidden layer. The input and hidden layers are connected by the activation function, $f = Sigmoid(W^{(1)} \cdot x + b^{(1)})$, in the mapping process, where $W^{(1)}$ is the weight matrix and $b^{(1)}$ is the bias vector. More specifically, each input vector x_i is mapped onto the hidden layer with the expression of significant, reduced features. In contrast, the decoder phase reconstructs the input layer x . Input data f is mapped back onto the output layer of \tilde{x} with high-dimensional reconstructed data. The activation function $\tilde{x} = Sigmoid(W^{(2)} \cdot f + b^{(2)})$ is used to connect the hidden layer to output layer, where the weight matrix $W^{(2)} = (W^{(1)})^T$ is interpreted as tied weights and $b^{(2)}$ is the bias vector of the decoder phase. The AE is optimized architecture with parameter sets $(W^{(1)}, W^{(2)}, b^{(1)}, b^{(2)})$ to minimize error of restructuring in the output layer. The following cost function is used:

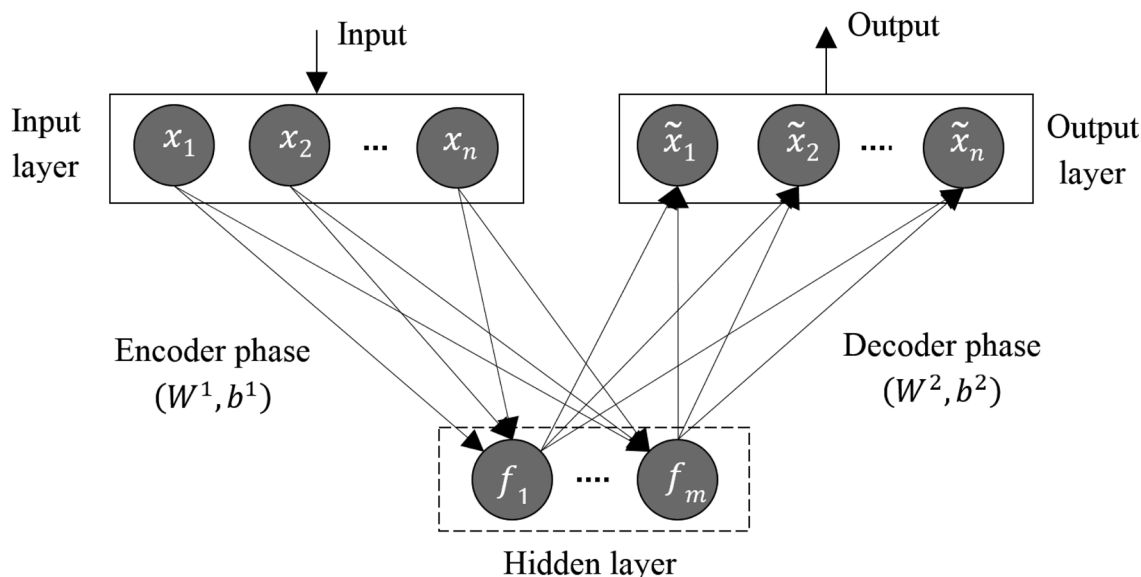


Fig. 1. Architecture of single AE network

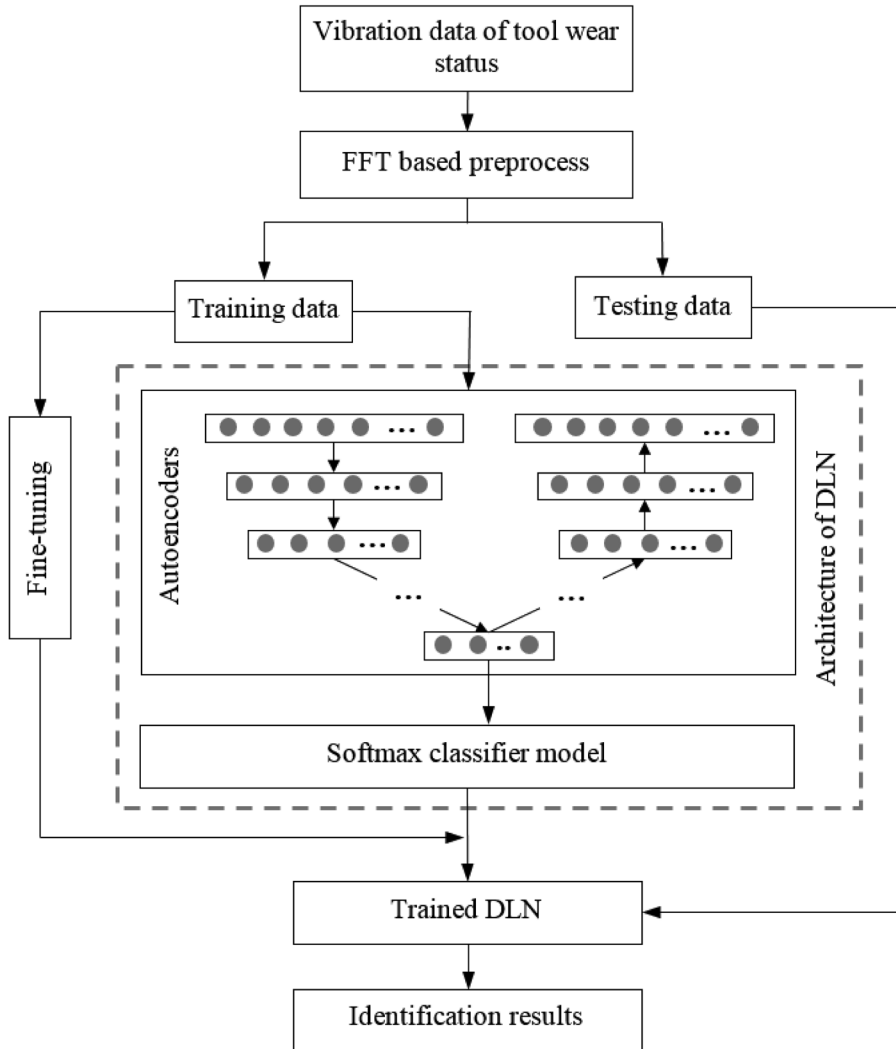


Fig. 2. Diagnosis technique based on DLN architecture

$$C = \frac{1}{N} \sum_n \sum_i (f_i^{(n)} - \tilde{f}_i^{(n)})^2 + \lambda \cdot \Omega_W + \beta \cdot \Omega_S, \quad (1)$$

where i is the number of variables in input data; N is the number of samples; λ is the coefficient for the Ω_W ; β is the coefficient for the Ω_S ; Ω_W is an L_2 regularization term defined by Eq. (2); and Ω_S is a sparsity regularization term defined by Eq. (3).

$$\Omega_W = \sum_l \sum_i \sum_j (W_{ij}^l)^2, \quad (2)$$

$$\Omega_S = \sum_l \sum_j KL(\rho \parallel p_j^k). \quad (3)$$

where p_j^k is the mean activation for unit j in layer k ; ρ is the desired mean activation; and KL is the Kullback-Leibler divergence, which is defined by Eq. (4).

$$KL(\rho \parallel p_j^k) = \rho \log \frac{\rho}{p_j^k} + (1 - \rho) \log \frac{1 - \rho}{1 - p_j^k}. \quad (4)$$

It can be seen that each AE is independently trained to represent the features. The features are extracted from hidden layer nodes that contain the most important information of the input layer. The extracted features can be input data for the next AE to produce higher-level features.

1.2 Softmax Classifier Model

As a follow-up, training the softmax classifier model to identify patterns is a necessary step for the whole model of the diagnosis technique. The model uses the encoded feature data in the hidden layer of the

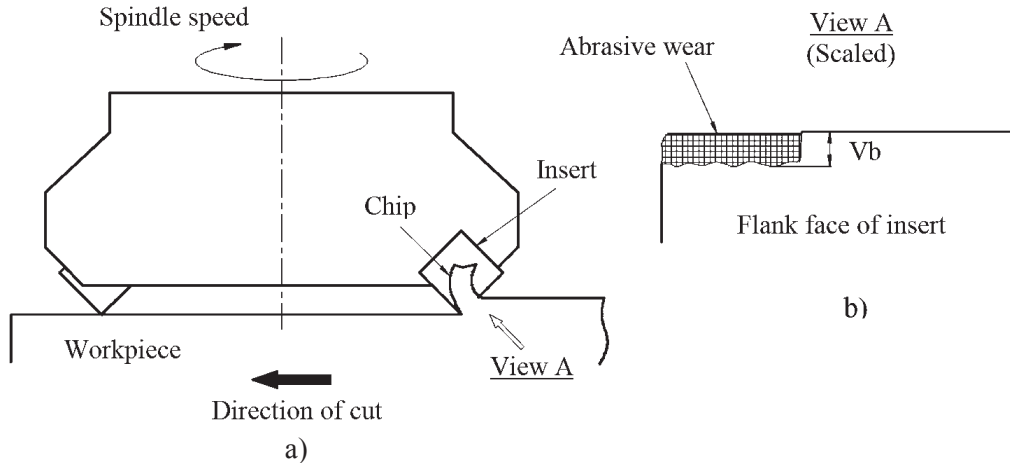


Fig. 3. Schematic drawing of face milling process; a) cutting zone at tool/workpiece interface, and b) observing of V_b tool wear on the insert

last AE. Softmax uses loss function based on cross entropy [16] and [23].

Softmax based on the supervised learning algorithm requires input samples $x = \{x^{(1)}, x^{(2)}, \dots, x^{(n)}\}$ and class labels $t = \{t^{(1)}, t^{(2)}, \dots, t^{(k)}\}$ for the classifier model. The training process for an input sample evaluates the probability: $probability(t=j|x)$ for each value of $j=1, 2, \dots, k$. This means that the probability of the class label is estimated by each of the k different possible values. Therefore, hypothesis function $\mathcal{H}_\theta(x^{(i)})$ is constituted as follows:

$$\mathcal{H}_\theta(x^{(i)}) = \begin{bmatrix} probability(t^{(i)} = 1|x^{(i)}; \theta) \\ probability(t^{(i)} = 2|x^{(i)}; \theta) \\ \vdots \\ probability(t^{(i)} = k|x^{(i)}; \theta) \end{bmatrix} = \frac{1}{\sum_{j=1}^k \exp(\theta_j^T x^{(i)})} \begin{bmatrix} \exp(\theta_1^T x^{(i)}) \\ \exp(\theta_2^T x^{(i)}) \\ \vdots \\ \exp(\theta_k^T x^{(i)}) \end{bmatrix}, \quad (5)$$

where $\theta_1, \theta_2, \dots, \theta_k \in R^{n+1}$ are the parameters of the softmax classifier, and the $\frac{1}{\sum_{j=1}^k \exp(\theta_j^T x^{(i)})}$ term represents the normalization of the distribution.

AEs and softmax classifier are hierarchically stacked together to construct the diagnosis technique. This technique is then fine-tuned with all the parameters to optimize the whole model, which will be evaluated using a test dataset.

2 PROPOSED DIAGNOSIS TECHNIQUE

This section describes the DLN in the proposed diagnosis technique. The significant feature data is extracted from vibration data related to tool wear status by unsupervised algorithm-based AEs, which provides the softmax classifier with inputs. The combination of AEs with softmax in the proposed DLN architecture achieves the impressive diagnosis results. Vibration data in the frequency domain corresponding to tool wear status are inputs to the DLN. Fig. 2 shows diagnosis implementation, which comprises the following seven steps:

- Step 1:* Acquire vibration data based on cutting tool status.
- Step 2:* Pre-process the collected data by FFT in time domain to express data clearly in the frequency domain.
- Step 3:* Divide the training data-testing ratio for the diagnostic phase. The training data are then used to train the model. The testing data are used to evaluate the trained model.
- Step 4:* Exploit the important features in the hidden layer AE, which is based on an unsupervised algorithm.
- Step 5:* Train the softmax classifier model using the extracted features of the last AE. Arrange the AEs and softmax classifier in the DLN architecture.
- Step 6:* Train the gained DLN by fine-tuning all the parameters using the class label.
- Step 7:* Identify the actual status of tool wear by the trained DLN.

3 EXPERIMENTAL RESULTS

This section presents the experimental results, demonstrating the quality of the proposed method. The wear status of face milling tool is identified in the face milling process. Vibration data on the state of the tool is collected, and pre-processed before serving as input to the proposed method.

3.1 Spindle Vibration Based Tool Wear Data Acquisition

The experimental dataset of BEST Lab at UC Berkeley is used to identify tool wear status [25]. The authors used a Matsuura machining center model MC-510V with a 70 mm face mill mounted with six inserts of the KC710 (Kennametal) type and a cast-iron milling workpiece is. The face milling operations dataset was experimented on different conditions, and the flank wear Vb values are verifiability measured, respectively. Fig. 3 is a schematic drawing of the face milling process. Vibration data was monitored by an accelerometer with a maximum sampling rate of 100 kHz. Table 1 shows a dataset of 40 samples. Samples were collected on machine spindle vibration corresponding to four tool wear states: $Vb_0=0$ mm, $Vb_1=0.11$ mm, $Vb_2=0.29$ mm, and $Vb_3=0.50$ mm,

which correspond to no tool wear, slightly worn tool, half worn tool, and severely worn tool, respectively. The acquired vibration data was recorded at the end of 2nd, 7th, 29th, and 44th minutes. Tool status was tested with the same cutting parameters at a spindle speed of 826 rpm, feed of 0.5 mm, and cutting depth of 1.5 mm. Fig. 4 shows these vibration signals in the time domain with four examples of corresponding wear statuses. The figure may imply that the vibration intensity that corresponds with cutting tool status is unclear and that the tool wear status cannot be determined even though the cutting parameter is unchanged.

Table 1. Tool wear statuses based vibration data

Cutting tool status	Training samples	Testing samples	Flank wear bandwidth [mm]	Class
Not worn Vb_0	(x_1-x_5)	(x_6-x_{10})	-	1
First worn Vb_1	$(x_{11}-x_{15})$	$(x_{16}-x_{20})$	0.11	2
Second worn Vb_2	$(x_{21}-x_{25})$	$(x_{26}-x_{30})$	0.29	3
Third worn Vb_3	$(x_{31}-x_{35})$	$(x_{36}-x_{40})$	0.50	4

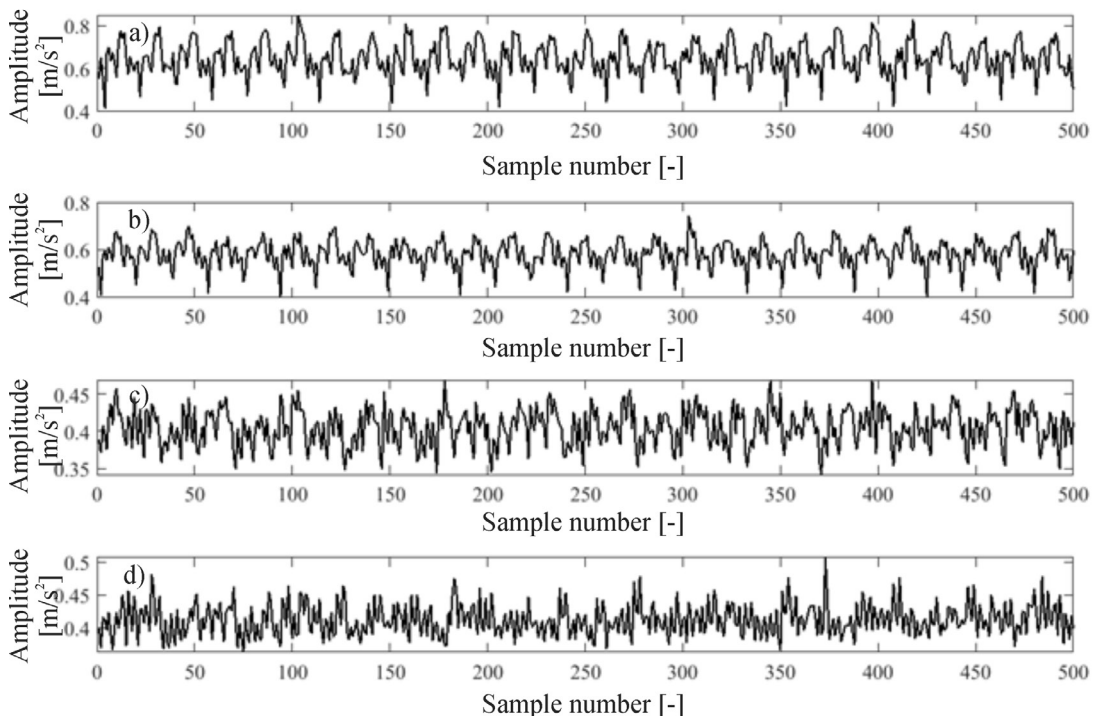


Fig. 4. Example of vibration signals use for cutting tool status determination; a) normal state; b) first-worn status; c) second-worn status, and d) third-worn status

3.2 Results and Discussion

To highlight the specific vibration frequencies of tool wear statuses, original time-domain vibration data was pre-processed and transformed into the frequency domain using the FFT method. These data were then used to extract the features of cutting tool status. The feature dataset of the vibration signal was extracted from two AEs, whose parameters are shown in Table 2.

Table 2. Parameter of AEs

	AE structure	λ	β
AE 1	512-20-512	0.05	6
AE 2	20-3-20	0.05	4

Table 3 shows the experimental results for identifying tool status. The results indicate each state of the cutting tool. The original vibration data were accurately identified by the class to which they belong. The diagnostic result for each tool wear status is 100 % accurate, which means that the expression of the unsupervised features of the original vibration signal exploited their efficient and important features. In this case, the combination of the optimized AE is related to Eq. (1) for the reconstruction of data with the softmax classifier model, which, according to the probability of Eq. (5) is very effective. Fig. 5 shows the confusion matrix for the four wear statuses of cutting tool (i.e., Vb_0 , Vb_1 , Vb_2 , and Vb_3), which are the same class as determined by the model trained by previous training data.

Table 3. Results of tool wear identification based on our proposed method

Tool wear status	Testing data	Features			Results
1	(x_6-x_{10})	0.559913	0.010705266	0.007923	5(1)
2	$(x_{16}-x_{20})$	0.055347	0.336805625	0.012599	5(2)
3	$(x_{26}-x_{30})$	0.000139	0.000127181	0.000102	5(3)
4	$(x_{36}-x_{40})$	0.000108	9.24E-05	8.32E-05	5(4)
Training time [s]					0.3072
Accuracy [%]					100

To compare the identification results with the other classifiers, the authors constructed the shallow classifiers as the feed-forward neural network (FNN) classifier and a k-nearest neighbour (k-NN) classifier to identify tool wear status. The extracted three-dimensional feature data of the last AE was used to train and test the classifiers. An FNN classifier with

10 hidden layer, four output layer, and training error goal of 0.01 and a k-NN classifier with four nearest neighbours are formed to conduct the identification procedure. Table 4 shows the evaluation result of the FNN classifier and the k-NN classifier. The evaluation showed that the identification accuracy of both these classifiers was lower than the identification accuracy of the proposed DLN. It is known that the tool wear diagnosis technique, based on the proposed DLN, effects high-level feature representation in deep learning to gain high classification accuracy. Nevertheless, the perfect classification accuracy results for the proposed DLN come at a high price in terms of time compared to the only k-NN architecture, as the results in Table 3 and 4 show. This causes each phase of DLN construction to be optimized. Finally, the DLN based our proposed diagnosis technique is confirmed as efficient for feature representation and classification, which is illustrated in Fig. 6.

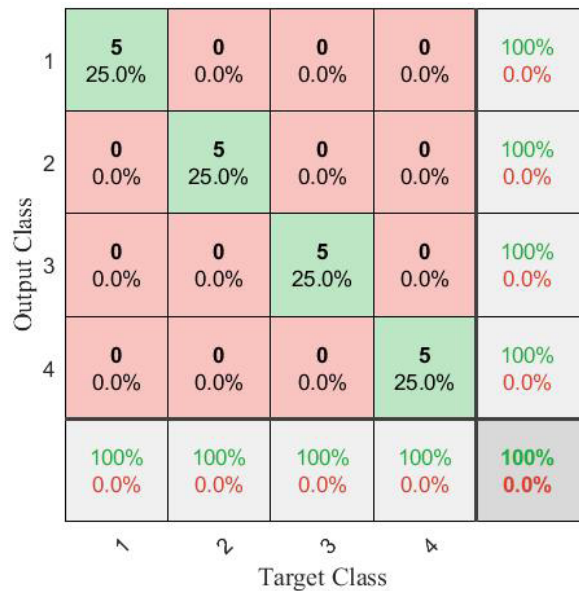


Fig. 5. Identification result confusion matrix

Table 4. Results of tool wear identification based on FNN and k-NN classifiers

Class	Testing data	FNN	k-NN
1	(x_6-x_{10})	5(1)	3(1)2(3)
2	$(x_{16}-x_{20})$	5(2)	5(2)
3	$(x_{26}-x_{30})$	0(3)5(4)	5(3)
4	$(x_{36}-x_{40})$	5(3)	0(4)5(3)
Time [s]		0.3462	0.1098
Accuracy [%]		75	65

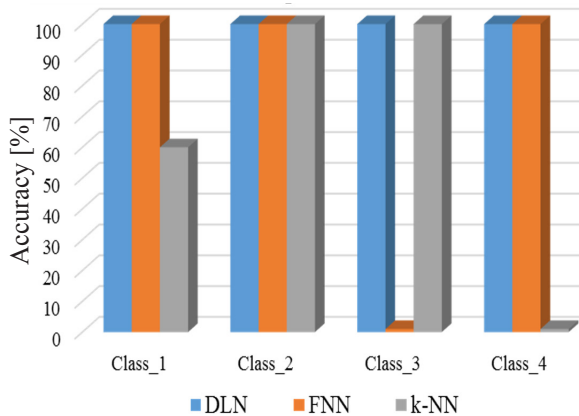


Fig. 6. Comparisons of identification accuracy results

4 CONCLUSIONS

This research, based on a theoretical framework and experiment, proposes a diagnostic method for determining the wear states of cutting tools used for face milling. The experimental data are the spindle vibration signals of the milling machine in relation to wear state of cutting tools. The proposed diagnosis technique is structured based on the SAE in feature representation, a softmax classifier model, and experimental data. In particular, the vibration data related the tool wear state are used by AEs to create compact low-dimensional data that expresses the most important features of the actual tool states. Softmax is then combined to train the model to recognize tool states. The experimental results demonstrate that this proposed technique is highly effective in accurately identifying patterns of tool wear states in the face milling process. Based on data of time series, the authors believe that this proposed technique can be used to diagnose other conditions of the cutting tool such as tool breakage or tool life and can be applied to other diagnostic fields such as the bearing fault or gear fault.

5 ACKNOWLEDGEMENT

The authors would like to express our gratitude to Faculty of Mechanical Engineering, Hanoi University of Industry.

6 CONFLICTS OF INTEREST

The authors declare that they have no conflict of interest.

7 REFERENCES

- [1] Lauro, C.H., Brandão, L.C., Baldo, D., Reis, R., Davim, J.P. (2014), Monitoring and processing signal applied in machining processes—A review. *Measurement*, vol. 58, p. 73-86, DOI:10.1016/j.measurement.2014.08.035.
- [2] Cui, Y., Fussell, B.K., Jerard, R.B., Esterling, D. (2009). Tool wear monitoring for milling by tracking cutting force model coefficients. *Transactions of the North American Manufacturing Research Institution of SME*, vol. 37, p. 613-620.
- [3] Hoang, D.T., Nguyen, N.-T., Tran, Q.D., Nguyen, T.V. (2019). Cutting forces and surface roughness in face-milling of SKD61 hard steel. *Strojniški vestnik - Journal of Mechanical Engineering*, vol. 65, no. 6, p. 375-385, DOI:10.5545/sv-jme.2019.6057.
- [4] Chi, Y., Dai, W., Lu, Z., Wang, M., Zhao, Y. (2018), Real-time estimation for cutting tool wear based on modal analysis of monitored signals. *Applied Sciences*, vol. 8, no. 5, art ID. 708, DOI:10.3390/app8050708.
- [5] Čerče, L., Pušavec, F., Kopač, J. (2015) A new approach to spatial tool wear analysis and monitoring. *Strojniški vestnik - Journal of Mechanical Engineering*, vol. 61, no. 9, p. 489-497, DOI:10.5545/sv-jme.2015.2512.
- [6] Lu, H., Chang, C., Hwang, N., Chung, C. (2009), Grey relational analysis coupled with principal component analysis for optimization design of the cutting parameters in high-speed end milling. *Journal of Materials Processing Technology*, vol. 209, no. 8, p. 3808-3817, DOI:10.1016/j.jmatprotec.2008.08.030.
- [7] Dimla, E.D. Snr. (2000) Sensor signals for tool-wear monitoring in metal cutting operations—a review of methods. *International Journal of Machine Tools and Manufacture*, vol. 40, no. 8, p. 1073-1098, DOI:10.1016/S0890-6955(99)00122-4.
- [8] Penedo, F., Haber, R.E., Gajate, A., Toro, R.M. (2012), Hybrid incremental modeling based on least squares and fuzzy K-NN for monitoring tool wear in turning processes. *IEEE Transactions on Industrial Informatics*, vol. 8, no. 4, p. 811-818, DOI:10.1109/TII.2012.2205699.
- [9] Brezak, D., Majetic, D., Udiljak, T., & Kasac, J. (2012) Tool wear estimation using an analytic fuzzy classifier and support vector machines. *Journal of Intelligent Manufacturing*, vol. 23, no. 3, p. 797-809, DOI:10.1007/s10845-010-0436-x.
- [10] Nguyen, V.H., Cheng, J.S., Thai, V.T. (2017). An integrated generalized discriminant analysis method and chemical reaction support vector machine model (GDA-CRSVM) for bearing fault diagnosis. *Advances in Production Engineering & Management*, vol. 12, no. 4, p. 321-336, DOI:10.14743/apem2017.4.261.
- [11] Palanisamy, P., Rajendran, I., Shanmugasundaram, S. (2008). Prediction of tool wear using regression and ANN models in end-milling operation. *The International Journal of Advanced Manufacturing Technology*, vol. 37, p. 29-41, DOI:10.1007/s00170-007-0948-5.
- [12] Li, X., Lim, B., Zhou, J., Huang, S., Phua, S., Shaw, K., Er. M.J. (2009). Fuzzy neural network modelling for tool wear estimation in dry milling operation. *Annual Conference of the Prognostics and Health Management Society*, p. 1-11.

- [13] Bajić, D., Celent, L., Jozić, S. (2012). Modeling of the influence of cutting parameters on the surface roughness, tool wear and the cutting force in face milling in off-line process control. *Strojniški vestnik - Journal of Mechanical Engineering*, vol. 58, no. 11, p. 673-682, DOI:10.5545/sv-jme.2012.456.
- [14] Zimroz, R., Bartelmus, W., Barszcz, T., Urbanek, J. (2014). Diagnostics of bearings in presence of strong operating conditions non-stationarity—A procedure of load-dependent features processing with application to wind turbine bearings. *Mechanical systems and signal processing*, vol. 46, no. 1, p. 16-27, DOI:10.1016/j.ymssp.2013.09.010.
- [15] Jia, F., Lei, Y., Lin, J., Zhou, X., Lu, N. (2016). Deep neural networks: A promising tool for fault characteristic mining and intelligent diagnosis of rotating machinery with massive data. *Mechanical Systems and Signal Processing*, vol. 72-73, p. 303-315, DOI:10.1016/j.ymssp.2015.10.025.
- [16] Bishop, C.M. (2006). *Pattern Recognition and Machine Learning*. Springer, Cambridge.
- [17] Fritscher, K., Raudaschl, P., Zaffino, P., Spadea, M.F., Sharp, G.C., Schubert, R. (2016). Deep neural networks for fast segmentation of 3D medical images. *Medical Image Computing and Computer-Assisted Intervention; 19th International Conference*, p. 158-165, DOI:10.1007/978-3-319-46723-8_19.
- [18] Gao, S., Zhang, Y., Jia, K., Lu, J., Zhang, Y. (2015). Single sample face recognition via learning deep supervised autoencoders. *IEEE Transactions on Information Forensics and Security*, vol. 10, no. 10, p. 2108-2118, DOI:10.1109/TIFS.2015.2446438.
- [19] Nguyen, V.H., Cheng, J.S., Yu, Y., Thai, V.T. (2019). An architecture of deep learning network based on ensemble empirical mode decomposition in precise identification of bearing vibration signal. *Journal of Mechanical Science and Technology*, vol. 33, no. 1, p. 41-50, DOI:10.1007/s12206-018-1205-6.
- [20] Chong, E., Han, C., Park, F.C. (2017). Deep learning networks for stock market analysis and prediction: Methodology, data representations, and case studies. *Expert Systems with Applications*, vol. 83, p. 187-205, DOI:10.1016/j.eswa.2017.04.030.
- [21] Lu, C., Wang, Z., Zhou, B. (2017). Intelligent fault diagnosis of rolling bearing using hierarchical convolutional network based health state classification. *Advanced Engineering Informatics*, vol. 32, p. 139-151, DOI:10.1016/j.aei.2017.02.005.
- [22] Långkvist, M., Karlsson, L., Loutfi, A. (2014). A review of unsupervised feature learning and deep learning for time-series modeling. *Pattern Recognition Letters*, vol. 42, p. 11-24, DOI:10.1016/j.patrec.2014.01.008.
- [23] Bengio, Y. (2009). Learning deep architectures for AI. *Foundations and Trends® in Machine Learning*, vol. 2, no. 1, p. 1-127, DOI:10.1561/2200000006.
- [24] Shi, C., Panoutsos, G., Luo, B., Liu, H., Li, B., Lin, X. (2019). Using multiple-feature-spaces-based deep learning for tool condition monitoring in ultraprecision manufacturing. *IEEE Transactions on Industrial Electronics*, vol. 66, no. 5, p. 3794-3803, DOI:10.1109/TIE.2018.2856193.
- [25] Agogino, A., Goebel, K. (2007). Milling data set. *NASA Ames Prognostics Data Repository*, NASA Ames, Moffett Field, <http://ti.arc.nasa.gov/project/prognostic-data-repository>.