Research on numerical control machine fault diagnosis based on distribution adaptive one-dimensional convolutional neural network

Zhang Yuying 1, Lu Yi 1,* and Zhao Jing 2

1College of Metrology and Measurement Engineering, China Jiliang University, Hangzhou, Zhejiang, China
2Hangzhou Wolei Intelligent Technology Co., Ltd, Hangzhou, 310018, China
*The corresponding author, E-mail: luyi9798paper@163.com

Abstract. Numerical control machine is a high-precision and high-automation equipment, if the problem occurred in operation, it can affect the processing conditions of the machinery parts first. If the fault is aggravated, it can eventually cause the numerical control machine to stop. Spindle bearings and tools are the most vulnerable parts of numerical control machine. Previously, resonance demodulation technique was used for bearing fault diagnosis. Empirical analysis or neural network was used for tool fault diagnosis. However, the numerical control machine is an entirety, the fault is usually caused by multi-dimensional factors, the above method doesn't work when two types of faults occur at the same time. To diagnose faults of numerical control machine, a fault diagnosis model named distribution adaptive deep convolutional neural networks (DADCNN) was proposed. This model was based on One-dimensional convolution algorithm. The Batch Normalization algorithm was involved to overcome the problem of changing data distribution in the middle layer. The t-SNE algorithm was used to visualize and view the feature classification results. Experiments show that the accuracy of this model can reach 90.29%, and it has good fault diagnosis capabilities.

1. Introduction
Numerical control machine is the main equipment of the manufacturing industry, it has the characteristics of high precision, high automation and high efficiency. Processing level of numerical control machine is directly related to many industries such as mechanical industry, aviation industry and automobile industry[1]. Spindle bearings is a key part of the numerical control machine. If the problem like abrasion and gluing occurred, the temperature of spindle bearings will rise under long-term use. If the fault is aggravated, it can affect the high-precision processing of the machinery parts, even it can cause the equipment down[2]. Based on the resonance demodulation technology, Liu Peisen diagnosed faults of spindle bearings by comparing the frequency appearing in the envelopment frequency spectra with the fault characteristic frequency[3]. The tool is in direct contact with the workpiece, and its wear state needs to be paid attention to. Tool changing in machining centers often depends on the experience level of the workers. If workers change tools too frequently, the manufacturing cost will increase. If workers do not change tools for a long time, the machining accuracy of the machinery part will decrease[4]. Dai Wen converted the tool vibration data into wavelet scale maps, and input them into a CNN model for training and diagnosis. The CNN model was based on Alexnet network. This method had a good effect on the judgment of tool wear status[4].
The above-mentioned fault diagnosis methods for spindle bearings and tools do not consider the condition that two types of faults occur at the same time. At this time, the vibration signal includes not only the characteristics of spindle bearings, but also the characteristics of tool wear. Therefore, to diagnose the faults of spindle bearings and tools, a fault diagnosis system for numerical control machine was proposed. This system collected the vibration signal of spindle bearings and tools, and input it into a one-dimensional convolutional network which contained Batch Normalization algorithm. This system can diagnose faults of spindle bearings and tools through learning.

2. General system design
The process of the numerical control machine fault diagnosis system is: First of all, the vibration acceleration sensor was used to collect the vibration signals of spindle bearings and tools. The data needed to be pre-processed to remove the influence of power frequency interference; Secondly, to simulate the faults of spindle bearing, fault signal obtained by simulation was added to the data. Then, a complete data set which contained different spindle bearing faults and different tool wear conditions could be obtained; Then this data set was used to train the deep learning model. The trained model was used to test and analyze the stage of numerical control machine. The overall design of the system is shown in Figure 1:

![Figure 1](image)

2.1. Hardware design
The hardware part of the system mainly includes numerical control machine, vibration acceleration sensor, constant-current source, data acquisition card and edge computer. The numerical control machine model is vmc-850. The vibration acceleration sensor model is 1A314E, which is used to collect vibration signal in the X, Y, and Z directions. The constant-current source model is YMC8204, which is used to output 10mA current to supply power to the sensors. The data acquisition card model is PCI-1716. The sampling frequency $f_s$ is 20kHz. The collected vibration signal is stored in the edge computer for subsequent model training. The schematic diagram of the hardware structure is shown in Figure 2:
2.2. Software design

The software part of the system includes data acquisition module, fault simulation module and fault diagnosis module. Data acquisition module is used to collect the vibration signal of numerical control machine through vibration acceleration sensor on the edge computer. The fault simulation module is used to simulate the fault signal of the spindle bearing. Since the experimental numerical control machine is an actual machine tool, the tool wear signal is easy to obtain, but the spindle bearing cannot be damaged deliberately. So the fault simulation signal is added as shown in equation (1), (2) [5,6]:

$$Y(t)=S(t)+r(t)$$  \hspace{1cm} (1)

$$S(t) = y_0 e^{-\zeta \omega t} \cos \omega \sqrt{1 - \zeta^2} t$$  \hspace{1cm} (2)

$S(t)$ is the periodic pulse attenuation signal, the amplitude is $y_0$, the damping coefficient is $\zeta$, the carrier frequency is $f_n$, the modulation frequency (fault characteristic frequency of spindle bearing) is $f_m = T^{-1}$, and $T$ is the period of the fault signal. $r(t)$ is the Gaussian white noise in the simulation signal, which is used to simulate the background noise in the simulation signal of bearing fault.

Set $y_0=2$, $\zeta=0.1$, $f_n=1100$Hz, $f_m=107.3$Hz, $r(t)$ is Gaussian white noise with a signal-to-noise ratio of 7dB. The time-domain waveforms of the fault simulation signal $Y(t)$ of the bearing outer ring and its two components are shown in Figure 3-Figure 5. The envelope spectrum of the fault simulation signal is shown in Figure 6. The comparison of the vibration signal before and after adding the fault simulation signal is shown in Figure 7 and Figure 8:

![Figure 2 The hardware structure schematic diagram.](image)

![Figure 3 The periodic pulse attenuation signal $S(t)$.](image)

![Figure 4 The white noise $r(t)$.](image)

![Figure 5 The fault simulation signal $Y(t)$.](image)

![Figure 6 The envelope spectrum of fault simulation signal.](image)
It can be seen from Figure 6 that a signal component with a frequency of 107.3Hz and its frequency multiplier is involved. This signal is the spindle bearing outer ring fault simulation signal of the numerical control machine. If this signal is added to the vibration signal, the vibration signal includes not only the characteristics of spindle bearing fault, but also the characteristics of tool wear. If other bearing fault simulation signal is needed, only need to modify the fault characteristic frequency of spindle bearing.

The fault diagnosis module uses the one-dimensional convolutional neural network to learn the vibration signal. The vibration signal is measured by the numerical control machine under different wear conditions of spindle bearing and tool. Then the learned model is used to diagnose and classify the test data to determine the state of the numerical control machine.

3. Fault diagnosis model based on distribution adaptive one-dimensional convolutional neural network

3.1. Algorithm description

Convolutional neural network often use stacked 3×3 convolution kernels in the field of deep learning. It can deepen the network depth and use fewer parameters to obtain a larger receptive field. Therefore, overfitting can be suppressed[7]. Based on this experience and combining the characteristics of one-dimensional signal, a model named Distribution Adaptive Deep Convolutional Neural Networks (DADCNN) was proposed to diagnose faults for numerical control machine. The first layer of the model uses a large-scale convolution kernel, whose role is to extract short-term features. The size of the large-scale convolution kernel is optimized by multiple experiments, and its function is to learn features related to fault diagnosis and prevent excessive learning due to noise interference. The remaining convolutional layers use 3×1 convolution kernels. The model proposed in this paper combines two convolution kernels of different scales, which guarantees the speed of calculation and suppresses the phenomenon of over-fitting under the premise of deepening the network depth.

In addition, the Batch Normalization algorithm was introduced for the problem of changing the data distribution in the middle layer during the training process. This algorithm can improve the generalization ability of the network and speed up the training speed, and it can also prevent the gradient from disappearing or exploding. The structure diagram of DADCNN is shown in Figure 9:
3.2. BN layer

In the training of the deep network, the nonlinear transformation of the network can change the data distribution of the middle layer. So the network needs to learn a new data distribution. Therefore, the continuous change of the data distribution during the training process causes the network's training speed to slow down. The Batch Normalization algorithm is proposed to overcome the covariance shift phenomenon and enhance the generalization ability of the network. The batch normalization layer is standardized for each mini-batch, its process is shown in equation (3)-(6). First, calculate the mean value \( \mu_B \) and variance \( \sigma_B^2 \) of the input \( x_i \) according to equation (3) (4). Second, standardize the input \( x_i \) according to formula (5). However, simple standardized operations restrict the data to a small range and limit the expressive ability of the model. Therefore, equation (6) introduces two parameters \( \gamma \) and \( \beta \) to scale and shift the standardized data to reconstruct the characteristic distribution of the data. The calculation steps of the BN layer are as follows:

\[
\begin{align*}
\mu_B &= \frac{1}{m} \sum_{i=1}^{m} x_i \\
\sigma_B^2 &= \frac{1}{m} \sum_{i=1}^{m} (x_i - \mu_B^2) \\
\hat{x}_i &= \frac{x_i - \mu_B}{\sqrt{\sigma_B^2 + \epsilon}} \\
y_i &= \gamma \hat{x}_i + \beta
\end{align*}
\]

where \( x_i \) is the i-th input in the mini-batch, \( m \) is the length of the mini-batch, \( \epsilon \) is a constant term to ensure numerical stability, \( \gamma \) is the scaling parameter, and \( \beta \) is the shifting parameter.

4. Fault diagnosis test

4.1. Test data preparation

The experimental data used in the research was collected in the vertical machining centers vmc-850, and the collection frequency was 20kHz. The experimental data is periodic. So a data enhancement method named overlap sampling was used to make full use of the experimental data. This method can also avoid over-fitting and enhance the generalization ability of the network.
As shown in Figure 10, the training sample length used in this article is 2048, so 2048 is used as the sampling window. If the length of a certain fault signal is 20000 and the offset is 1, 17953 training samples can be made.

The numerical control machine faults involved in this test are classified into 8 categories. That is, the fault-free bearings, outer ring fault bearings, inner ring fault bearings and rolling element fault bearings corresponding to the scrap tool and the new tool respectively. The specific classification labels are shown in Table 1:

| Data set classification label | new tool | scrap tool |
|-------------------------------|----------|------------|
| fault-free bearings           | 0        | 1          |
| outer ring fault bearings     | 2        | 3          |
| inner ring fault bearings     | 4        | 5          |
| rolling element fault bearings| 6        | 7          |

### 4.2. Test model and parameters

The model parameters used in the experiment are shown in Table 2. The model has 5 convolutional layers. The first layer uses a large-scale convolution kernel to extract features, and the remaining layers use a smaller convolution kernel to increase the depth. The Softmax layer has 8 outputs, corresponding to 8 categories.

| Structural parameters of the DADCNN model | number | Network layer | Convolution kernel size/step size | Number of convolution kernels | Output size (width × depth) | spot patch |
|------------------------------------------|--------|---------------|----------------------------------|-------------------------------|-----------------------------|------------|
| 1                                        | Convolutional layer 1 | 128×1/16×1 | 8                                | 128×8                         | Yes                         |
| 2                                        | Pooling layer 1       | 2×1/2×1    | 8                                | 64×8                          | No                          |
| 3                                        | Convolutional layer 2 | 3×1/1×1    | 16                               | 64×16                         | Yes                         |
| 4                                        | Pooling layer 2       | 2×1/2×1    | 16                               | 32×16                         | No                          |
| 5                                        | Convolutional layer 3 | 3×1/1×1    | 16                               | 32×16                         | Yes                         |
| 6                                        | Pooling layer 3       | 2×1/2×1    | 16                               | 16×16                         | No                          |
| 7                                        | Convolutional layer 4 | 3×1/1×1    | 32                               | 16×32                         | Yes                         |
| 8                                        | Pooling layer 4       | 2×1/2×1    | 32                               | 8×32                          | No                          |
| 9                                        | Convolutional layer 5 | 3×1/1×1    | 32                               | 6×32                          | Yes                         |
| 10                                       | Pooling layer 5       | 2×1/2×1    | 32                               | 3×32                          | No                          |
| 11                                       | Fully connected layer | 64          | 1                                | 64×1                          | No                          |
| 12                                       | Softmax layer         | 8           | 1                                | 8                             | No                          |
4.3. Test model training
The DADCNN model and the model without BN layers are used for experimentation. The curves that the accuracy and loss value of the two models vary with the training times is shown in Figure 11 and Figure 12:

![Figure 11](image1.png)  **Accuracy curve.**

![Figure 12](image2.png)  **Loss curve.**

It can be concluded from the above two figures that the accuracy of the DADCNN model has increased rapidly compared with the model without BN layer, and the training process has less fluctuation. The model without BN layer takes more time and fluctuates greatly during the training process. Its accuracy rate is not as good as the DADCNN model. It shows that the model without BN layer has a general convergence effect. Therefore, the Batch Normalization algorithm has great advantages in improving the stability of model training and shortening the training time.

4.4. Fault diagnosis result
In order to verify the stability of the DADCNN model, 20 experiments were carried out. The diagnosis accuracy and loss value of each test varies with the experiment times as shown in Figure 13:

![Figure 13](image3.png)  **Diagnosis accuracy and loss value change curve with the number of experiments.**

In 20 experiments, the diagnostic accuracy of the DADCNN model is between 88.08%-92.17%. The average accuracy is 90.29%. The loss value fluctuates between 0.22 and 0.36, with an average value of 0.28, which proves that the model has good learning and diagnostic capabilities.

5. Model visualization analysis
t-SNE is a non-linear dimensionality reduction algorithm, which is suitable for dimensionality reduction of high-dimensional data to 2 or 3 dimensions for visualization. In this experiment, the t-SNE algorithm is used to visualize the feature expression of each convolutional layer. The classification results can be
viewed as an important basis for network parameter adjustment. All convolutional layers’ feature expression result produced by the t-SNE algorithm is shown in Figure 14 - Figure 18. Labels 0-7 correspond to the types of faults in Table 1. It can be seen that after 5 times of convolution, 8 types of fault features can be distinguished basically, which shows that the model has good classification ability.

Figure 14   The first convolutional layer.

Figure 15   The second convolutional layer.

Figure 16   The third convolutional layer.

Figure 17   The fourth convolutional layer.

Figure 18   The fifth convolutional layer.
6. Conclusion
The DADCNN model has a good effect on the fault diagnosis of numerical control machine. Compared with the traditional CNN model, this model uses the Batch Normalization algorithm to speed up training and improve accuracy. The t-SNE algorithm was used to visualize and view the feature classification results. However, it can be seen from the experimental results that this model still has misjudgments, this problem will be solved in the future.

Acknowledgment
This work was financially supported by key research and development project of Zhejiang Province (2021C01136).

References
[1] Xin Mei. Application of neural network training algorithm in fault diagnosis of NC machine tool [J]. Techniques of Automation and Applications, 2020,39(02):13-17.
[2] Wen Yan, Tan Jiwen and Li Shan. Fault diagnosis for rolling bearing based on multiple classifiers fusion and fuzzy [J]. China Sciencepaper, 2016,11(04),464-469
[3] Liu Peisen. Research on fault diagnosis instrument of rolling bearing based on the theory of resonance demodulation [D]. University of Electronic Science and Technology of China, 2014.
[4] Dai Wen, Zhang Chaoyong, Meng Leilei, Xue Yanshe, Xiao Pengfei and Yin Yong. Prediction model of milling cutter wear status based on deep learning [J]. China Mechanical Engineering, 2020,31(17),2071-2078
[5] Yan Xiaoa and Jia Minping. Morphological demodulation method based on improved singular spectrum decomposition and its application in rolling bearing fault diagnosis [J]. Journal of Mechanical Engineering, 2017,53(07),104-112
[6] Li bing, Zhang Peilin, Liu Dongsheng, Mi Shuangshan and Ren Guoquan. Feature extraction for roller bearing fault diagnosis based on adaptive multi-scale morphological gradient transformation [J]. Journal of Vibration and Shock, 2011,30(10):104-108
[7] Zhang Wei, Li Chuanhao, Peng Gaoliang, Chen Yuanhang and Zhang Zhujun. A deep convolutional neural network with new training methods for bearing fault diagnosis under noisy environment and different working load[J]. Mechanical Systems and Signal Processing, 2018, 100:439-453