CNN4GCDD: a One-Dimensional Convolutional Neural Network-based Model for Gear Crack Depth Diagnosis

Shouhua Zhang
Faculty of Information Technology and Electrical Engineering
University of Oulu
Oulu, Finland
shouhua.zhang@oulu.fi

Jiehan Zhou
Faculty of Information Technology and Electrical Engineering
University of Oulu
Oulu, Finland
jiehan.zhou@ieee.org

Erhua Wang
School of Intelligent Equipment
Changzhou College of Information Technology
Changzhou, China
ehuaw@126.com

Susanna Pirttikangas
Faculty of Information Technology and Electrical Engineering
University of Oulu
Oulu, Finland
susanna.pirttikangas@oulu.fi

***Abstract*—Gear crack is one of the common failures in transmission systems. With the gradual expansion of cracks, it may cause tooth fracture. Therefore, it is of great significance to study the fault diagnosis of gear cracks. Vibration signals with time sequence are widely used in gear fault diagnosis. Extracting key fault features from vibration signals determines the accuracy of fault diagnosis models. This paper takes spur gears as research objects, and proposes a model for diagnosing gear crack depth based on one-dimensional convolutional neural network (short for CNN4GCDD). In order to identify crack depths, we collect the vibration signals from three gears with various crack depths and a normal gear without cracks. CNN4GCDD uses the original vibration signal as the input, adaptively extracts features, and makes crack depth diagnosis through the convolutional neural network. The experimental results demonstrate that CNN4GCDD can directly use the original time-domain signal for crack depth diagnosis, and make a high accurate prediction.**

Keywords—convolution neural network, deep learning, fault diagnosis, gear crack depth

I. INTRODUCTION

Gears are important variable speed and transmission components in various types of machinery. Gears are prone to partial failures due to poor working conditions. Gear failure affects the operating status of the entire equipment if it cannot be checked on time. A study found that 65% of gearbox damage is due to gear faults [1]. The detection and diagnosis of gear status has always been a research hotspot in the field of rotating machinery fault diagnosis. Among the forms of gear failures, tooth fracture has the greatest impact on the gearbox, which often results in the entire gearbox being scrapped. Machines may be damaged in severe cases due to excessive instantaneous impact. It is inconvenient to carry out experimental verification due to the sudden and instantaneous hazard. Therefore, the early manifestation of gear tooth fracture is of great significance to the fault diagnosis of gear crack.

At present, the use of vibration signals for fault diagnosis is still one of the mainstream methods in the fault pattern recognition of gears [2]. It is possible to make accurate judgments on the operating status of equipment by extracting and analyzing the characteristics of vibration signals. The main causes of gear vibration are pitch line impact and meshing impact during transmission. When a gear is in a normal or abnormal state, the meshing frequency vibration component and its harmonics always exist, but the vibration levels of the two states are different. This is also the theoretical basis for diagnosing gear faults through vibration signals.

The traditional fault diagnosis based on vibration signal needs to analyze the internal operation mechanism of the mechanical system first, and then use signal processing to analyze fault signals. This process needs to rely on professional knowledge, and the quality of the extracted features directly affects the effect of fault diagnosis.

Compared with traditional machine learning, deep learning has more powerful feature extraction and processing capabilities and has been widely used in various fields, such as image classification [3][4], face detection [5][6], natural language processing [7][8], etc. Methods based on deep learning can adaptively extract features from vibration signals, including not only well-known fault features, but also some potential features that can be used to identify faults and are difficult to be defined and explained explicitly. Therefore, more and more researches apply deep learning into gear fault diagnosis (GFD).

Wang et al. proposed an adaptive normalized convolutional neural network (CNN) to diagnose different fault locations and severities of planetary gearbox in the scenarios of complex variable working conditions and data imbalance accurately and automatically [9]. Liu et al. trained a CNN using singular value decomposition matrices as inputs and achieved planetary GFD [10]. Shi et al. constructed a novel deep neural network (DNN) based on bidirectional-convolutional long short-term memory (LSTM) networks to determine the type, location, and direction of planetary gearbox faults by extracting spatial and temporal features from both vibration and rotational speed measurements automatically and simultaneously [11]. Yin et al. proposed a fault diagnosis method for wind turbine gearboxes based on optimized LSTM with cosine loss [12]. Tao et al. proposed a multilayer gate recurrent unit (GRU) method for spur GFD [13].

There are still few studies on the diagnosis of gear crack depth (GCD) although a large number of studies have used deep learning theory to study GFD. Extracting key fault features from one-dimensional (1D) vibration signals can improve the accuracy of fault diagnosis models. 1D-CNN performs extremely well in analyzing time sequence sensor data [14][15]. 1D-CNN can extract local features like other CNNs. Features it learns at one location can be recognized at other locations. Therefore, this paper proposes a simple and efficient end-to-end one-dimensional CNN model (i.e., CNN4GCDD) which directly takes one-dimensional raw data as input, extracts features, classify and diagnoses GCD. The structure of the model is relatively simple, easy to train, and can well ensure the stability of the accuracy of the test set.

The remainder of the paper is organized as follows. Section II reviews one-dimensional CNN. Section III presents the model for diagnosing gear crack depth based on one-dimensional convolutional neural network (short for CNN4GCDD) and a multi-layer LSTM model for comparison. Section IV shows the experiments and result analysis. Section V concludes the paper and presents the future work.

II. ONE-DIMENSIONAL CNN

CNN is a deep feedforward neural network with powerful feature extraction capabilities. It constructs multiple filters that can extract features of the input data and uses these filters to extract the representative features hidden in the input data layer by layer. At the same time, it combines sparse connections and parameter weight sharing mechanisms to reduce dimensionality and sampling precision in time and space, reduce the data dimension, reduce the amount of training parameters, and effectively avoid algorithm overfitting.

1D-CNN performs a dot product operation on the data of the convolution window to extract local features. The features it learns at a certain position can be recognized at other positions due to its translation invariance. 1D-CNN also has pooling operations, using MaxPooling1D and AveragePooling1D to complete maximum pooling and average pooling. 1D-CNN can be well applied to time sequence analysis of sensor data, such as gyroscope or accelerometer data. It can also be well used to analyze signal data with a fixed length period, such as audio signals. In addition, it can also be applied to natural language processing. 1D-CNN is mainly composed of one-dimensional convolution layer, one-dimensional pooling layer, full connection layer and classifier, as shown in Figure 1 [16].

1D convolution layer: 1D convolution layer is composed of many 1D convolution kernels. The main function of convolution kernels is to learn the feature representation of input data. 1D convolution only convolves in one direction. The output of 1D convolution of the i-th convolution kernel is.

$$y_i = f(\sum X \Theta k_i + b_i) \;\; i \in K \tag{1}$$

where f is an activation function, ⊝ stands for the convolution, k_i is the i-th convolution kernel, b_i is a bias, and K is the number of convolution kernels and the number of the channels. An activation function can be sigmoid, tanh or relu.

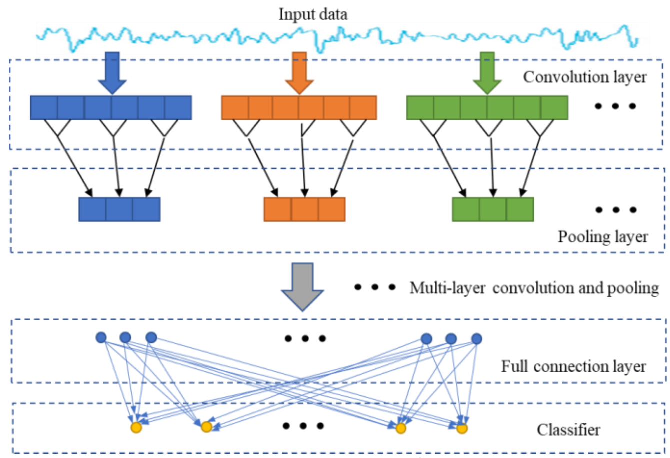

Fig.1. The structure of 1D-CNN

1D pooling layer: The role of pooling layers is mainly to reduce the dimensionality of the feature vector output by the convolutional layer, and to improve the result at the same time, so that the structure is not prone to overfitting. Through convolutional layers and pooling layers, more abstract features can be obtained. The output of the i-th channel of l-length after pooling is

$$P_i(j) = \max\big(y_i(j * W, (j+1) * W)\big) \;\; 0 \le j \le l/s \tag{2}$$

where W is the width of the pooling window, and s is the stride size.

Full connection layer: The output of the full connection layer is

$$\delta = f(\omega P + b) \tag{3}$$

where ω is weight, b is bias, and f is an activation function.

Classifier: The classifier uses the softmax activation function.

III. DIAGNOSIS OF GEAR CRACK DEPTH

A. CNN4GCDD

One-dimensional vibration signals are widely used in gear fault diagnosis, so that the gear transmission system can be maintained in time to reduce losses. Extracting key fault features

TABLE I. STRUCTURAL PARAMETERS OF CNN4GCDD

Network layer	Kernel	Channels
Convolution 1	width=64, stride=1	16
MaxPooling 1	width=2, stride=2	16
Convolution 2	width=4, stride=1	64
MaxPooling 2	width=2, stride=2	64
Convolution 3	width=4, stride=1	256
MaxPooling 3	width=2, stride=2	256
Convolution 4	width=4, stride=1	256
MaxPooling 4	width=2, stride=2	256
Convolution 5	width=4, stride=1	512
MaxPooling 5	width=2, stride=2	512
Convolugtion 6	width=2, stride=1	512
MaxPooling 6	width=2, stride=2	512

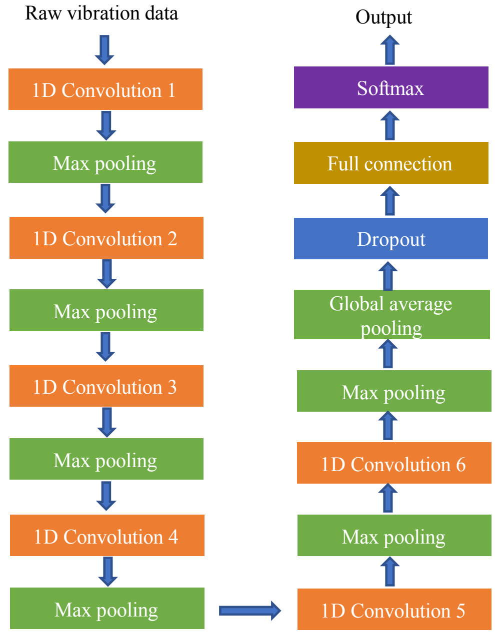

Fig. 2. The structure of CNN4GCDD

from vibration signals determines the accuracy of GFD. The effect of 1D-CNN in vibration signal analysis is comparable to that of recurrent neural networks (RNN), and the computational cost and parameter amount are much smaller.

This paper proposes a diagnosis model (CNN4GCDD) based on 1D-CNN for intelligent diagnosis of GCD. Figure 2 presents the structure of CNN4GCDD. There are 6 1D convolution layers, 6 1D pooling layers, 1 global average pooling layer and 1 dropout layer in the model. Table 1 presents the parameters used in the model. CNN4GCDD uses multi-layer 1D-CNN as the feature extractor and introduces global average pooling and dropout before the full connection layer, reducing the amount of trainable parameters and testing time. CNN4GCDD does not require any manual feature extraction and feature transformation operations on the original data during the entire fault diagnosis process. It only needs to input the original fault data into the model, and the fault diagnosis results are automatically output.

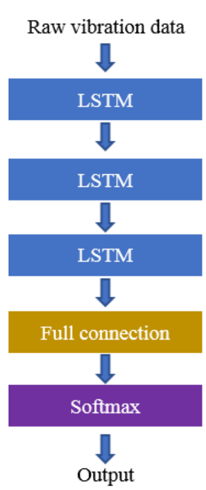

Fig 3. The structure of the multi-layer LSTM model

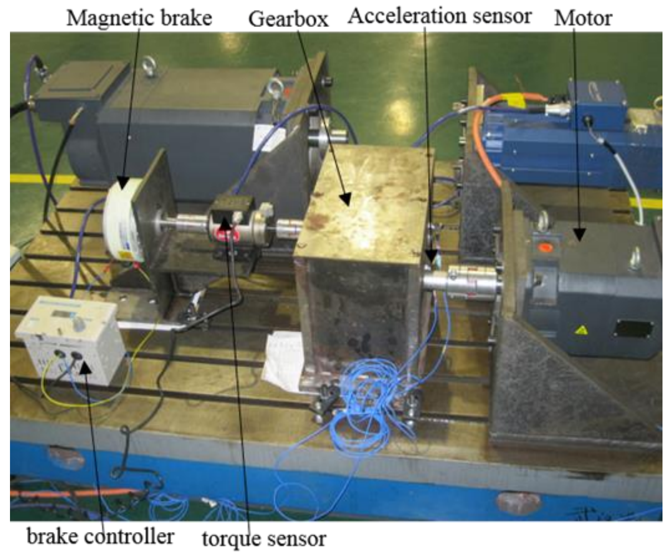

Fig 4. The experimental platform

B. *Multi-layer LSTM*

RNN is one of the most commonly used models when dealing with time series problems using deep learning. The reason why RNN has excellent performance on time series data is that RNN takes the output of hidden nodes of time slice t-1 as input at time slice t. The information of the previous time slice is also used to calculate the content of the current time slice, while the output of hidden nodes of a traditional model only depends on the input features of the current time slice.

LSTM is a special kind of RNN that solves the problem of long-term dependencies, and vanishing and exploding gradients during training with long sequences. Simply, LSTM can perform better in longer sequences than ordinary RNNs.

LSTM networks can not only deal with long-term dependencies of time series data, but also effectively deal with nonlinear and non-stationary problems of signals. Therefore, LSTM is widely applied in fault diagnosis based on vibration signals. In order to compare with CNN4GCDD, this paper also constructs a multi-layer LSTM model. Figure 3 presents its structure. The dimensions of the output space of the first layer, second layer and third layer LSTM are 64, 128, 512, respectively.

IV. EXPERIMENTS AND RESULT ANALYSIS

We use a first-stage reduction gearbox as the condition monitoring experiment platform, which includes a servo motor, a first-stage reduction gearbox, a three-dimensional acceleration sensor, a torque sensor, a magnetic brake and a brake controller, as shown in Figure 4. The acceleration sensor model PCB-356A16 is installed on the input shaft of the experimental platform. Table 2 shows the parameters of the driving and driven gears in the gearbox.

TABLE II. PARAMETERS OF THE GEARS

Type	Number of teeth	Module	Tooth width
Driving gear	50	2mm	20mm
Driven gear	80	2mm	20mm

Fig. 5. Gears with and without cracks

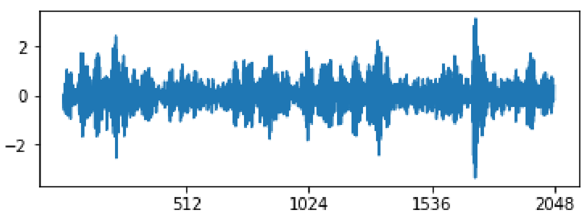

Fig.6-(a) C0

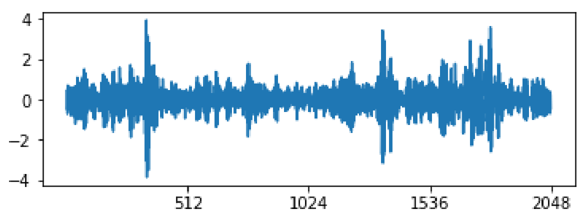

Fig. 6-(b) C1

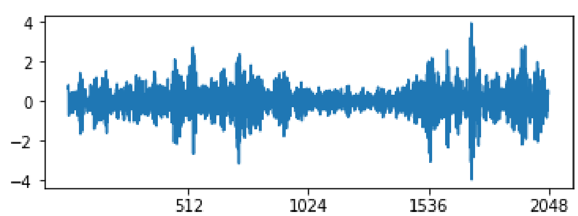

Fig. 6-(c) C2

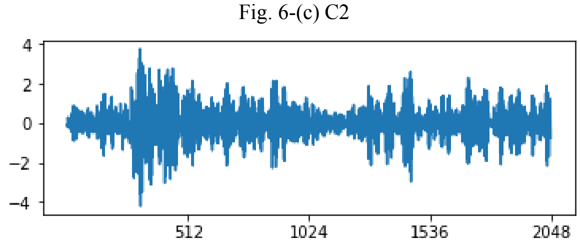

Fig.6-(d) C3

Fig.6. Time domain signals with 4 different conditions

TABLE III. PERFORMANCE COMPARISON BETWEEN CNN4GCDD AND MULTI-LAYER LSTM

Points	**Speed**	**Accuracy**	
		CNN4GCDD	*Muti-layer LSTM*
2048	300rpm	100%	41.86%
	600rpm	100%	40.00%
	900rpm	100%	50.00%
	300,600,900rpm	94.70%	60.43%
1024	300rpm	97.67%	43.02%
	600rpm	100%	49.45%
	900rpm	100%	49.44%
	300,600,900rpm	89.06%	71.69%
512	300rpm	94.19%	41.28%
	600rpm	100%	42.31%
	900rpm	99.44%	52.81%
	300,600,900rpm	70.43%	79.70%

The driving gears have radial cracks of different lengths including no cracks, ¼ cracks, ½ cracks and ¾ cracks, as shown in Figure 5. The crack length is $L_i = i\times(R_c\text{-}r)/4$, $I = 0, 1, 2, 3$. R_c is the radius of the root circle and r is the radius of the spindle hole. Radial cracks are processed by wire cutting. The acceleration signal of the input shaft is collected through NI PXI-1042, the load is 6 N·m, the rotating speeds are 300rpm, 600rpm and 900rpm respectively, and the sampling sample and got a training set of 1057 samples and a test set of 266 samples. Figure 6 presents the time domain signals in 4 different conditions at a speed of 600 rpm and a load of 6 N·m.

The four types of faults with gear crack lengths of 0, 1/4, 1/2, and 3/4 are expressed as C0, C1, C2, and C3, respectively. We took 512, 1024, 2048 points as an input sample respectively in experiments. The proportions of the training set and test set are 80% and 20%, respectively. Table 3 shows the experimental results on the test set. The accuracy of CNN4GCDD is much higher than that of the multi-layer LSTM model although the number of trainable parameters of the constructed multi-layer LSTM is larger than that of CNN4GCDD. Taking 2048 points as a sample works best for CNN4GCDD. The accuracy on a single-speed data set is higher than that on a multi-speed data set for CNN4GCDD. It is the exact opposite for the LSTM model.

V. CONCLUSION

GFD plays an important role in equipment maintenance. This paper proposes a simple and efficient end-to-end CNN model, that is CNN4GCDD, which directly takes vibration signals as input and can make the fault classification relatively simpler, easier to train. The experimental results demonstrate that CNN4GCDD can use the original time-domain signal for crack depth diagnosis and make a higher accurate prediction than LSTM method. Future work will focus on improving and testing the model in different working conditions to further improve the generalization ability of the model.

ACKNOWLEDGMENT

The authors operate under Academy of Finland 6Genesis Flagship program (grants 318927) SRA4 Services and Applications.

REFERENCES

[1] A. Aherwar, "An investigation on gearbox fault detection using vibration analysis techniques: A review," *Australian Journal of Mechanical Engineering*, vol.10, no.2, pp.169-183,2012.

[2] Z. K. Abdul, A. K. Al-Talabani and D. O. Ramadan, "A hybrid temporal feature for gear fault diagnosis using the long short term memory," IEEE Sensors Journal, vol. 20, no. 23, pp. 14444-14452, 2020.

[3] S. Li, W. Song, L. Fang, Y. Chen, P. Ghamisi and J. A. Benediktsson, "Deep learning for hyperspectral image classification: an overview,|" *IEEE Transactions on Geoscience and Remote Sensing*, vol. 57, no. 9, pp. 6690-6709, 2019.

[4] J. Liu et al., "Applications of deep learning to MRI images: a survey," *Big Data Mining and Analytics*, vol. 1, no. 1, pp. 1-18, 2018.

[5] G. Guo and N. Zhang, "A survey on deep learning based face recognition," *Computer Vision and Image Understanding*, vol. 189, no.2019, pp. 102805, 2019.

[6] M.Wang and W. Deng, "Deep face recognition: a survey," *Neurocomputing*, vol. 429, pp. 215-244, 2021.

[7] D. W. Otter, J. R. Medina and J. K. Kalita, "A Survey of the Usages of Deep Learning for Natural Language Processing," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 32, no. 2, pp. 604-624, 2021.

[8] I. Lauriola, A. Lavelli and F. Aiolli, "An introduction to Deep Learning in Natural Language Processing: Models, techniques, and tools," *Neurocomputing*, vol. 470, pp. 443-456, 2022.

[9] C. Wang, H. Li, K. Zhang, S. Hu and B. Sun, "Intelligent fault diagnosis of planetary gearbox based on adaptive normalized CNN under complex variable working conditions and data imbalance," Measurement, vol. 180, pp. 109565, 2021.

[10] D. Iatsenko, P. McClintock and A. Stefanovska, "Nonlinear mode decomposition: A noise-robust, adaptive decomposition method," *Phys. Rev. E*, vol.92, no. 3, pp.032916, 2015.

[11] J. Shi et al, "Planetary gearbox fault diagnosis using bidirectional-convolutional LSTM networks," *Mechanical Systems and Signal Processing*, vol. 162, pp. 107996, 2021.

[12] A. Yin, Y. Yan, Z. Zhang Z, C. Li and R. V. Sánchez, "Fault diagnosis of wind turbine gearbox based on the optimized LSTM neural network with cosine loss," *Sensors*, vol. 20, no. 8, pp. 2339, 2020.

[13] Y. Tao, X. Wang, R. Sánchez, S. Yang and Y. Bai, "Spur gear fault diagnosis using a multilayer gated recurrent unit approach with vibration signal," *IEEE Access*, vol. 7, pp. 56880-56889, 2019.

[14] O. Abdeljaber, O. Avci, S. Kiranyaz, M. Gabbouj and D. J. Inman, "Real-time vibration-based structural damage detection using one-dimensional convolutional neural networks," *Journal of Sound and Vibration*, vol. 388, pp. 154-170, 2017.

[15] W. Zhang, C. Li, G. Peng, Y. Chen and Z Zhang, "A deep convolutional neural network with new training methods for bearing fault diagnosis under noisy environment and different working load," *Mechanical Systems and Signal Processing*, vol. 100, pp. 439-453, 2018.

[16] J. Yu and X. Zhou, "One-dimensional residual convolutional autoencoder based feature learning for gearbox fault diagnosis," *IEEE Transactions on Industrial Informatics*, vol. 16, no. 10, pp. 6347-6358, 2020.