Sound Visualization and Convolutional Neural Network in Fault Diagnosis of Electric Motorbike

Jian-Da Wu*, Che-Yuan Hsieh, Wen-Jun Luo

Graduate Institute of Vehicle Engineering, National Changhua University of Education, Changhua, Taiwan

Corresponding Author Email: jdwu@cc.ncue.edu.tw

ABSTRACT

This study proposed convolutional neural network (CNN) training for different figure recognition to diagnose electric motorbike faults. Traditional motorbike maintenance is usually carried out by technicians to find the problem step by step. Many resources are wasted and time consumed in diagnosing maintenance problems. Due to rising environmental protection awareness, motorbike power systems gradually transformed from combustion engines into the electric motor. The sound amplitude generated by the combustion engine is great and may cover other faulty sounds. The electric power system sound amplitude is greatly decreased, permitting various fault diagnosis to be performed by extracting the electric motor sound. With the development of computers and image processing, deep learning neural network for picture recognition technology becomes more feasible. This study presents the motor system sound visualization for fault diagnosis. First obtain the sound signals of the motor in the five different states of the operation in the laboratory and the road test, and draw the time domain graph, frequency domain graph and spectrogram graph to be used as the test database. The results graphs of various states were trained through a CNN. The signal states were then classified to achieve fault diagnosis. Experiments and identification results show that the spectrogram and CNN method can identify motorbike faults most effectively compared to the time domain graph and the frequency domain graph.

1. INTRODUCTION

The motorbike is an indispensable means of transportation for most people in Taiwan. Once the motorbike is damaged, it may cause inconvenience in life and may endanger personal safety. Therefore, it is necessary to find a feasible and effective method to assess the status of the motorbike. After the power of the motorbike is gradually changed from the traditional engine to the current motor, the sound signal from the engine has been greatly reduced. Since the sound signal from the motor is very weak, we can record the sound signal of the entire electric motorbike. First, the fast Fourier transform (FFT) is used for feature extraction, and then the sound signal is recognized through the convolutional neural network (CNN) to diagnose the fault of the electric motorbike.

The concept of voice recognition was proposed before the computer was invented. As early as 1920, a toy company launched a toy dog called Radio Rex. When the dog's name was called, the spring inside would remove it from the base pops up [1], the earliest system that uses computers for voice recognition is the Audrey speech recognition system developed by Davis et al. [2] in 1952. Audrey can only easily recognize the pronunciation of the numbers 0-9, and the system achieved an accuracy rate of 98%. Its identification method is to track the formants in speech, and its accuracy reaches the level of being able to identify the speaker. In 1959, Denese of the College of London added grammatical probability to speech recognition [3]. Among them, the biggest breakthrough in speech recognition is the first large-vocabulary speech recognition system Sphinx based on the hidden Markov model proposed by Leonard in 1988, replacing the original method of guessing only by text or sound characteristics [4].

Since the computer was invented in the 1950s, scientists have hoped that computers can be used to develop artificial intelligence (AI). The term AI was first organized by Minsky in Dartmouth in 1961 [5]. Participants included several contemporary scholars. Among them, Newell and Simon showed logic theorist [6], which is known as "the world's first AI program". This is a program that automatically proves the theorem. However, due to the low hardware performance at that time and the lack of data, AI did not develop smoothly. It was not until the 1980s that the term AI was brought up again. At this time, scientists used a large number of statistical theories such as probability and statistics to allow computers to learn through data, also known as machine learning (ML) [7]. There are many different models of ML, and artificial neural network (ANN) [8] was the most popular when it first appeared. However, the hardware at that time had not yet reached the function of being able to be calculated immediately, the storage space was also small, and the amount of data was not enough, so it encountered obstacles, and the support vector machine (SVM) [9] has become the mainstream model at the time. Until the continuous improvement of technology and computer hardware, in 2006, Hinton et al. [10] successfully solved the problems encountered by ANN, and the ANN were replaced with the name "deep learning" [11]. Today, deep learning technology have a profound impact on

all major industries.

The first real CNN was proposed by LeCun et al. [12] in 1989. CNN method is the most widely used among many neural networks, and they have shown very good results on many machine vision problems. In addition, they have been successfully applied in the fields of computer graphics and natural language processing. In 2016, Janssens et al. [13] proposed a state monitoring feature learning model through the CNN. The feature learning method of convolution is compared with traditional feature engineering. This result shows that the feature learning system accumulated in convolution is better than the traditional feature engineering method.

In 2020, a study of malicious webpage detection algorithm based on image semantics that derived by the backpropagation neural network (BPNN) to obtain the target image, and combine the image with other functions of the malicious website, so that the malicious website can be detected [14]. In 2019, Fenjiro and Benbrahim proposed an optimal combination of imitation and reinforcement learning for self-driving cars, carry out self-driving training through deep learning [15]. In 2019, a micro-expression recognition algorithm for students in classroom based on CNN to detect human faces and detect facial micro-expressions is studied [16]. In 2019, an improved crowd counting method based on scale-adaptive CNN to obtain high-quality crowd density estimation maps, which are used to improve crowd statistics techniques was proposed [17]. In 2018, Wen et al. [18] proposed an effective method for automatically extracting the features of raw data. The signal is converted into a two-dimensional image through a conversion method, and the CNN is used to diagnose the motor bearing faults. Comparing the results with other methods, the results show that the method of CNN for bearing fault diagnosis has obvious effect. In 2018, Monteiro et al. [19] used Fourier transform spectrograms and CNN to analyze the vibration signals provided by the accelerometer to classify the faults of the gearbox. The results show that it is a competitive method. The self-driving car [20] is a product of the rapid development of ML. The self-driving technology has made rapid progress in recent years. The power part has also changed from the previous engine to the current motor [21]. Motors can increase operating efficiency and reduce energy consumption. The motor as a component of electrical energy conversion kinetic energy, whether it is used in vehicles, industrial, medical, etc. have the trace of the motor. In this study, four sets of different faults were simulated and a set of normal operating conditions were tested in laboratory and road test. The sound signal emitted by the electric motorbike is obtained through the sound measurement system, and then the picture of the signal is obtained through the analysis and drawing functions. After preprocessing the pictures, the result graphs of different states are then used to create a database, and then the CNN is used for learning and training. The obtained results can be used as identification of different states.

2. PRINCIPLE OF FEATURES EXTRACTION AND CONVOLUTIONAL NEURAL NETWORK

2.1 Principle of fast Fourier transform

When analyzing the sound signal, the physical quantity of the sound is expressed by the magnitude of the voltage, which is usually in the time domain. After fast Fourier transform (FFT), the time domain signal becomes a superposition of different sine wave signals. Analyzing the frequency of these sine waves can transform the signal into the frequency domain. Some signals are difficult to see in the time domain, but after converting to the frequency domain, it is easy to see the features. This is why many signal analysis uses FFT conversion. In addition, the FFT can extract the spectrum of a signal, which is often used in spectrum analysis. The application of FFT is not limited to digital signal processing. In fact, the FFT can speed up the training process of CNN. FFT is a discrete Fourier transform (DFT) that quickly calculates a sequence. It is used because the calculation of DFT is too complicated, and FFT can reduce this disadvantage and greatly reduce the amount of calculation. The DFT algorithm is described in the equation.

Let X_k, …,X_{N-1} be complex numbers. The DFT is shown as

$$X_k = \sum_{n=0}^{N-1} x_n e^{\frac{-i2\pi kn}{N}} \ (0 \le \mathrm{k} \le \mathrm{N}-1) \tag{1}$$

where, $e^{\frac{i2\pi}{N}}$ is a primitive N*th* root of 1. Direct evaluation according to this definition requires O (N^2) operations: X_k has a total of N outputs, and each output requires N sums. Direct use of DFT operation requires the use of N complex multiplications ($4N$ real number multiplications) and N-1 complex number additions ($2N$-2 real number additions), therefore, calculation of all N points using DFT requires N^2complex multiplication and N^2- N complex addition. FFT is a method that can calculate the same result in O ($N \ log \ N$) operations. Figure 1 shows how FFT transform time domain into frequency domain.
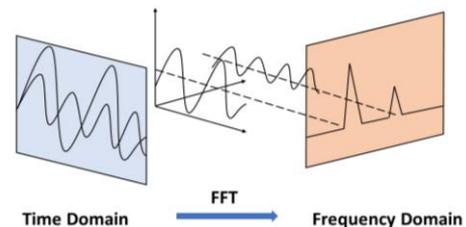

Figure 1. Principle of fast Fourier transform

2.2 Principle of convolutional neural network

The CNN structure was actually proposed very early. In 1989, LeCun proposed the idea of using backward propagation to train the convolutional layer [22]. However, it did not attract widespread attention at the time. The real modern CNN was proposed in 1998, LeCun proposed a new architecture after re-examining the CNN for nearly ten years. This new architecture was later known as the LeNet-5 CNN architecture [23]. In this new version, LeCun introduced advanced concepts such as shared weights, local connection and pooling compression, which laid the foundation for the contemporary CNN. Because of the LeNet-5 proposal enough knowledge was provided so that Krizhevsky et al. [24] proposed AlexNet which finally completely transformed the neural network into contemporary deep learning in 2012.

The CNN structure includes the input layer, the convolutional layer, the pooling layer and the fully connected layer as shown in Figure 2. The main convolutional neural network architecture is shown in Figure 3 consisting of seven blocks, with four convolutional layers and four pooling layers, finally leading to a fully connected layer. The input layer is the picture to be recognized by the input system. Whenever CNN recognizes a new picture, CNN will compare every place in the picture without knowing where the above features are. The mathematical principle behind this mechanism is called convolution, which is where the name CNN comes from. The convolutional layer is the core part of the CNN, usually composed of tens to hundreds of $n \times n$ filters. Each filter will enhance different image patterns as shown in Figure 4(a). The actual enhanced image mode of these filters is also found by the training process, so the convolutional layer can produce different filter effects for different problems. The convolution process is shown as [25].
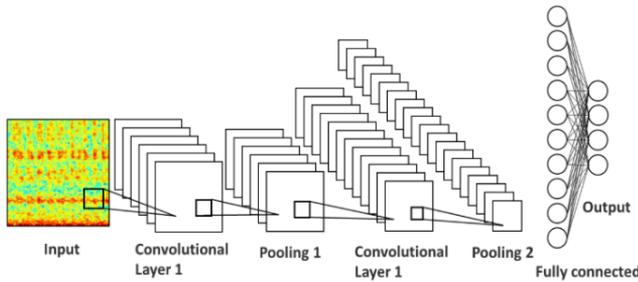

Figure 2. Principle of convolutional neural network

$$x_j^l = f\left(\sum_{i \in M_j} x_i^{l-1} * k_{ij}^l + b_j^l\right) \quad (2)$$

where, $(*)$ represents the convolution operation; M_j represents a selection of input maps; l is the lth layer in the network; k is the kernel matrix with size S×S; f is a nonlinear activation function.

The pooling layer is similar to down sampling in signal processing, and is usually connected after the convolution layer. After the original image is pooled, the number of pixels it contains will be reduced to one-fourth of the original ad shown in Figure 4(b). The CNN commonly used for image recognition will have one to three convolutional layers added to the pool when processing input data. The chemical layer processing is followed by two or more fully connected layers to output the prediction result. The pooling process is shown in Eq. (3) [25].

$$x_j^l = f\left(\beta_j^l \, down\left(x_j^{l-1} + b_j^l\right)\right) \quad (3)$$

where, down(.) represents a sub-sampling function. Typically, this function sums over each distinct n-by-n block in the input image so that the output image is n-times smaller along both spatial dimensions. Each output map is given its own multiplicative bias β and additive bias b. The fully connected layer will collect advanced filtered images and convert the information of these functions into weights, as shown in Figure 4(c). When input a picture to this machine, it will treat all pixel values as a one-dimensional list instead of the previous two-dimensional matrix. All values for voting for different options will be expressed in weight or connection

strength. Whenever CNN judges a new picture, the picture goes through a lot of low-level processing before it reaches the fully connected layer. After voting, the option with the highest number of votes will become the category of the picture.
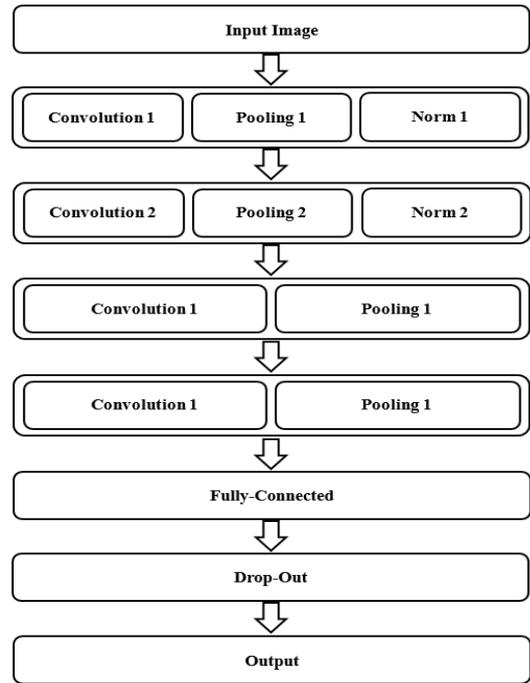
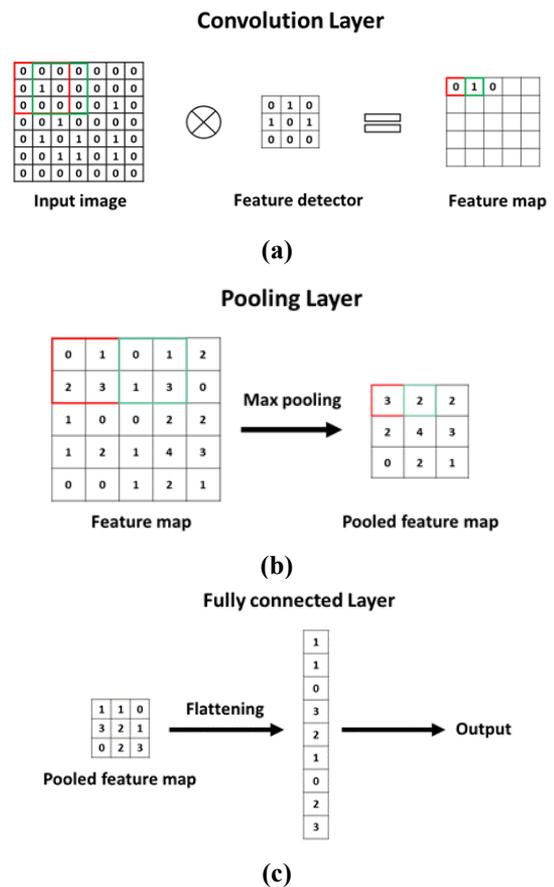

Figure 3. Convolutional neural network model

Figure 4. Convolutional neural network. (a) Principle of convolution layer, (b) Pooling layer, (c) Principle of fully connect layer

3. IMPLEMENTATION AND EXPERIMENTAL WORK

3.1 Experimental work and signal processing

This research will distinguish the various faults under different sounds in the experiment. Sound measurement will be divided into laboratory measurement and road test measurement. In this experiment, the sounds of KYMCO CANDY 2.0 EV electric motorbike were captured. In order to be able to recognize different states, five different faults were designed, A. Normal condition, B. Improper installation of the mudguard, C. Improper adjustment of the rear drum brake, D. Unlocked rear wheel nut, E. Insufficient lubrication of the gear box, which are summarized in Table 1.

In the laboratory test, the equipment used to collect audio files includes a microphone (PCB 426E01), a data acquisition system (NI-6024E), a data acquisition card (NI-9233), and a meter (RM-1500). The car was fixed onto a laboratory platform. A microphone was placed onto a windproof and sound-absorbing sponge. The sponge was inserted 10 cm above the electric motorbike motor. The motorbike conversion was controlled by a conversion meter. The alternating faults respectively captured five different alternative sound signals. There are 25 different conditions at 15 km/h, 20 km/h, 25 km/h, 30 km/h, 35 km/h, and repetitively capture 60 sound signals for each speed, each sound signal is captured for 1 second, and the sampling frequency is 4 kHz, as shown in Figure 5. After the sound is retrieved by the microphone, the analog signal received by the data record retrieval card is converted to a digital signal and stored in the computer using software.

Meanwhile, the road experiment was selected on an asphalt road near a campus. The professional equipment ZOOM H4n handheld digital recorder was used on the experimental equipment. The recorder was fixed above the electric motorbike and the actual sound status was recorded on the road. There are five different speed sound signals: 15 km/h, 20 km/h, 25 km/h, 30 km/h, 35 km/h. There are 25 different conditions, and at the same time, 60 sound signals were repeatedly acquired each time. The audio signal is captured for 1 second and the sampling frequency is 44.1 kHz built into the recorder.

Table 1. Specification of electrical vehicles in the experiment

Number	Status
A	Normal condition
B	Improper installation of the mudguard
C	Improper adjustment of the rear drum brake
D	Unlocked rear wheel nut
E	Insufficient lubrication of the gearbox

Figure 5. Experimental setup and process of test

3.2 Signal processing

This experiment simulates five different faults of electric motorbike at five different speeds, and is divided into laboratory data and road test data. The collected sound signals are drawn into a picture every second through fast Fourier transformation through a computer. The pictures are divided into time domain, frequency domain, and spectrogram graphs. Each fault records a 60-second sound signal at each speed. The flowchart of experimental process and fault diagnosis for electric motorbike is shown in Figure 6. After the sound signal is converted into a picture, the picture will be output as 128×128. After the picture is preprocessed, the CNN method is used for training and classification to obtain the recognition result.
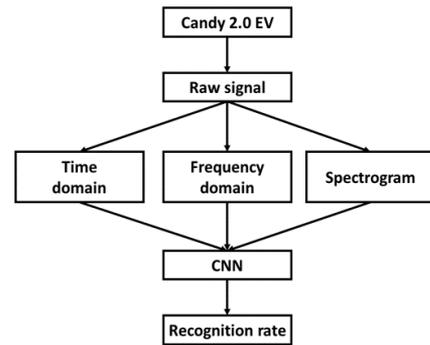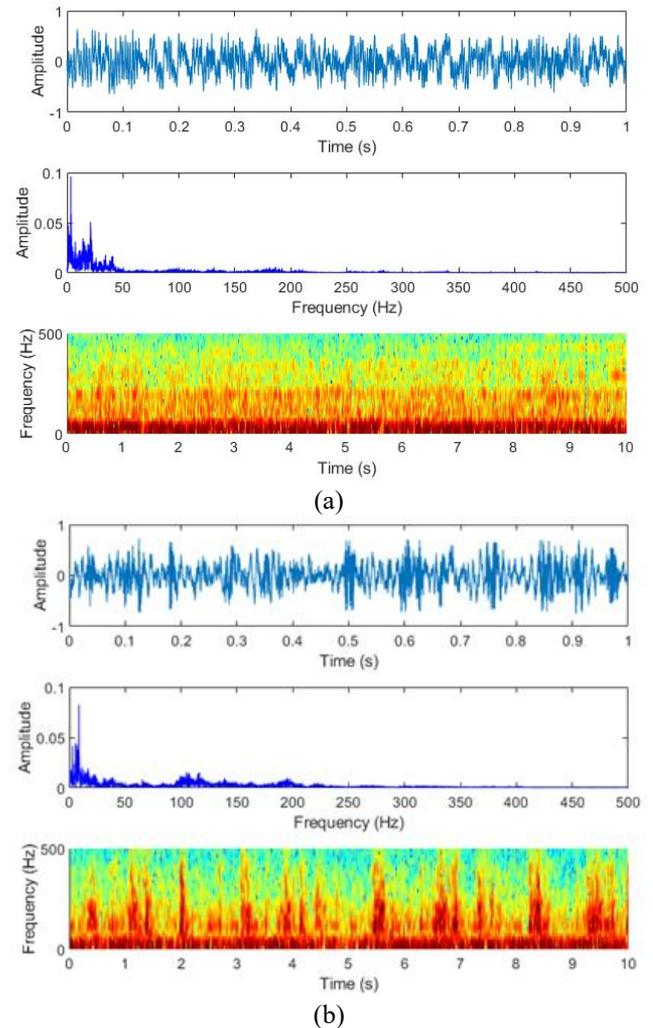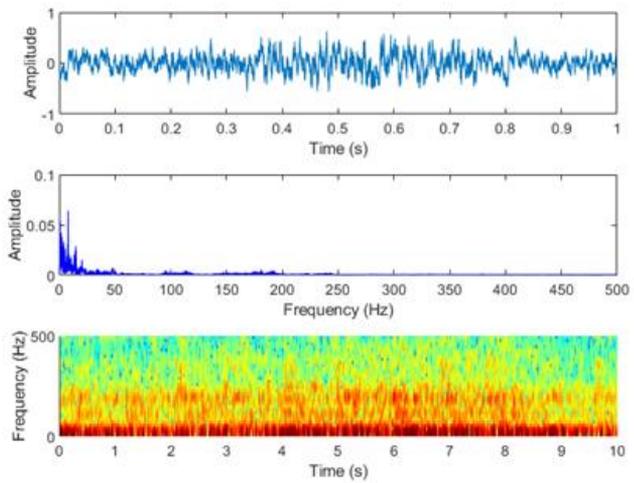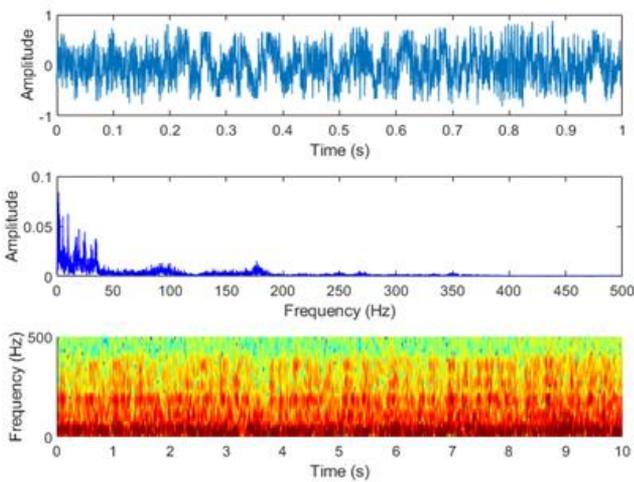

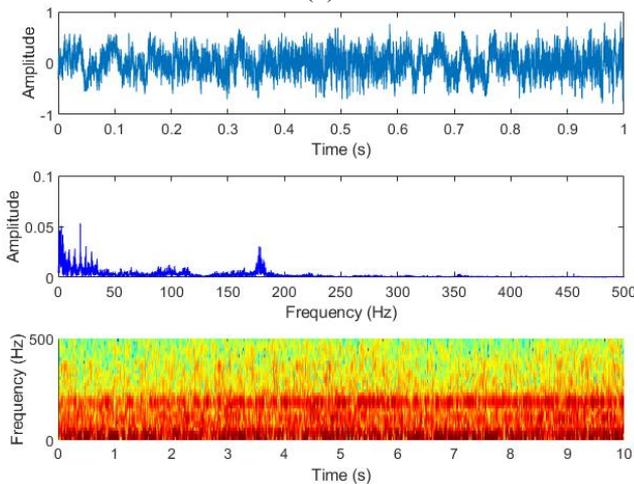**Figure 6.** Fault diagnosis flow chart for electric motorbike

(a)

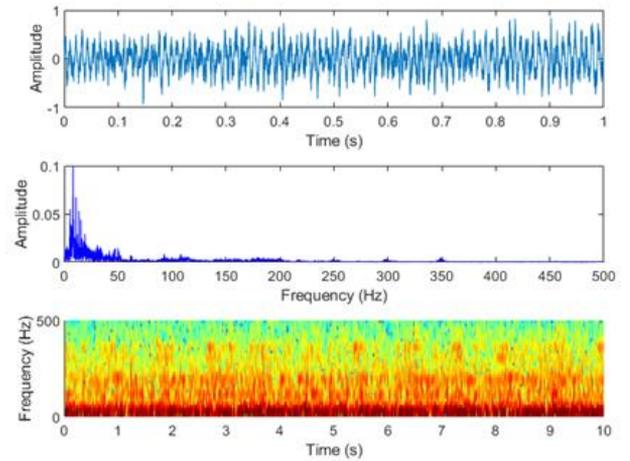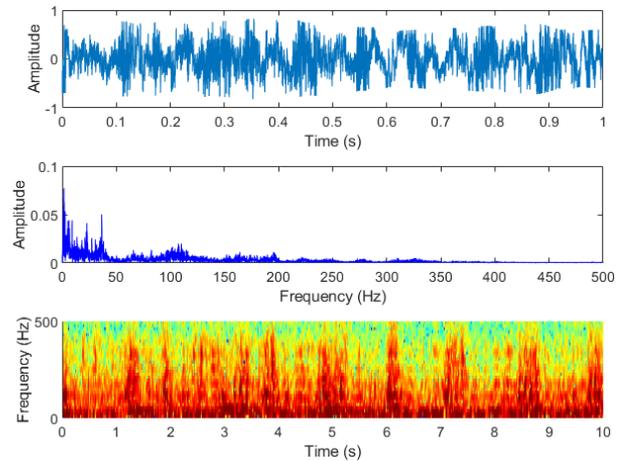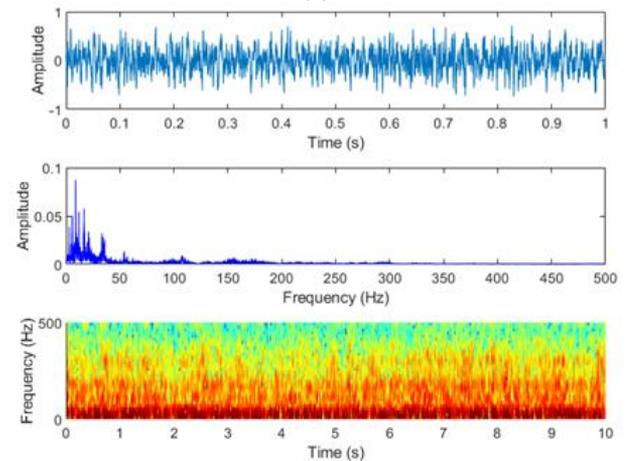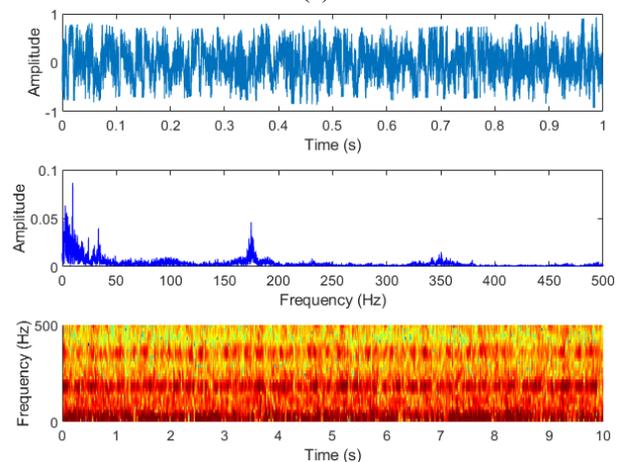
(b)

(c)

(d)

(e)

Figure 7. Laboratory test of electric motorbike fault sound visualization in five situations. (a) Normal condition, (b) Improper installation of the mudguard, (c) Improper adjustment of the rear drum brake, (d) Unlocked rear wheel nut, (e) Insufficient lubrication of the gear box

Figure 7 shows the laboratory test of electric motorbike fault sound visualization in five situations: (a) Normal condition (b) Improper installation of the mudguard (c) Improper adjustment of the rear drum brake (d) Unlocked rear wheel nut (e) Insufficient lubrication of the gear box. Figure 8 shows the road test of electric motorbike fault sound visualization in five situations.
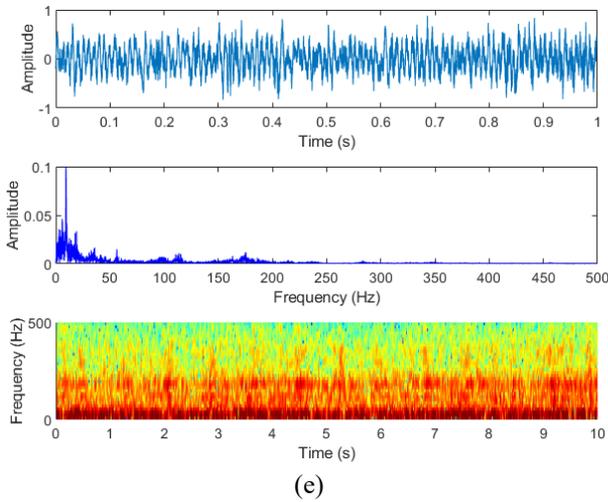

(a)

(b)

(c)

(d)

1823

Figure 8. Road test electric motorbike fault sound visualization in five situations. (a) Normal condition, (b) Improper installation of the mudguard, (c) Improper adjustment of the rear drum brake, (d) Unlocked rear wheel nut, (e) Insufficient lubrication of the gear box

4. EXPERIMENTAL RESULTS AND DISCUSSION

In the signal processing, this research assists the completion of the research through the use of the Anaconda package with the Spyder compiler, NumPy and Pandas suitable for data processing, the Matplotlib package that can visualize data, and Scikit-Learn, which provides a large number of algorithms and many practical data sets. Three kinds of pictures captured using different features are classified and identified by the CNN. From the results, it can be found that five conditions at five speeds are identified and the relationship between the number of samples and the diagnosis classification. Table 2 shows the laboratory test recognition rate of three feature extraction pictures for fault diagnosis at different layers at 15 km/h. In the identification process, using CNN in different sample numbers from 10 to 300 for three different special patterns. From the identification process, the data recognition rate of the time domain is the lowest. When the number of samples reaches 300, the recognition rate is only 80%. When the number of samples is 150, the recognition rate can reach 96.67% in frequency domain. The recognition rate of 100% can be reached when the number of training samples is 250. In the spectrogram data picture, the recognition rate is 100% when the number of training samples is 50. Table 3 shows the laboratory test recognition rate for three feature extraction pictures for fault diagnosis at different layers at 30 km/h. The experimental results also show similar results, and spectrogram data picture has the best recognition rate. Table 4 shows the road test recognition rate for three feature extraction pictures for fault diagnosis at different layers at 15 km/h. Table 5 shows the road test recognition rate for three feature extraction pictures for fault diagnosis at different layers at 30 km/h. Five condition include normal condition, improper installation of the mudguard, improper adjustment of the rear drum brake, unlocked rear wheel nut and insufficient lubrication of the gear box.

The time domain performance of the original signal is the lowest in the laboratory or road test data. The recognition rate only rises slowly, even if the recognition rate can reach 100% in the later period. The number of training layers needed is far more than the other two pictures. The frequency domain has undergone FFT extraction, and the recognition rate is only achieved at 150 layers. The recognition rate is slightly better than the original signal in the time domain. The spectrogram recognition rate with energy added is the highest among the three feature extraction pictures. The recognition rate can reach 100% when the training layer reaches 50 layers.

Figure 9 shows the laboratory test recognition rate for three feature extraction pictures for fault diagnosis at different layers at various speeds. Figure 10 shows the road test recognition rate for three feature extraction pictures for fault diagnosis at different layers at various speeds. Obviously, the laboratory environment is much quieter than road test in outdoors. In the laboratory test, the environmental noise impact on the signal is greatly reduced. Oppositely, in the road test, due to the increased wind noise and the friction between the tires and ground and other external factors, the sound captured is much more complicated. When using CNN for training and recognition, noise has an important impact on CNN identification. Compared with laboratory test, the recognition rate is relatively reduced. The graph shows that the road test data does not rise steadily like the laboratory test, but fluctuates. It is also speculated that due to the increase in external factors, the sound signal from the electric motor is relatively small, which affects the recognition rate. Although the road test results cannot be as good as the laboratory data, in actual conditions, the road test data can still have a very good recognition rate if the number of training samples is enough. For example, the Spectrogram graph is at a speed of 30 km/h. h, can reach 100% recognition rate.

Table 2. Laboratory test recognition rate for three feature extraction pictures for fault diagnosis at different layers at 15 km/h

Features \ Samples	Time Domain	Frequency domain	Spectrogram
10	16.66%	23.33%	26.66%
20	16.66%	30%	46.66%
30	23.33%	36.66%	50%
40	36.66%	53.33%	73.33%
50	23.33%	63.33%	100%
100	36.66%	73.33%	100%
150	63.33%	96.66%	100%
200	63.33%	96.66%	100%
250	66.66%	100%	100%
300	80%	100%	100%

Table 3. Laboratory test recognition rate for three feature extraction pictures for fault diagnosis at different layers at 30 km/h

Feature \ Samples	Time Domain	Frequency domain	Spectrogram
10	10%	13.33%	16.66%
20	20%	20%	26.66%
30	33.33%	33.33%	56.66%
40	40%	56.66%	86.66%
50	40%	100%	100%
100	60%	100%	100%
150	73.33%	100%	100%
200	76.66%	100%	100%
250	100%	100%	100%
300	100%	100%	100%

Table 4. Road test recognition rate for three feature extraction pictures for fault diagnosis at different layers at 15 km/h

Features Samples	Time Domain	Frequency domain	Spectrogram
10	10%	10%	46.66%
20	16.66%	16.66%	50%
30	16.66%	16.66%	56.66%
40	23.33%	20%	63.33%
50	30%	20%	73.33%
100	40%	36.66%	86.66%
150	50%	46.66%	93.33%
200	53.33%	76.66%	93.33%
250	56.66%	86.66%	96.66%
300	73.3%	100%	100%

Table 5. Road test recognition rate for three feature extraction pictures for fault diagnosis at different layers at 30 km/h

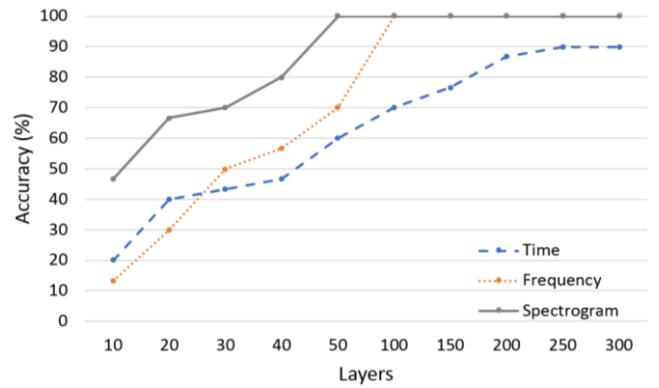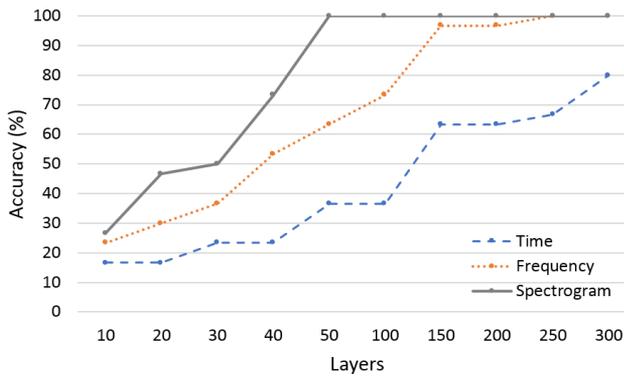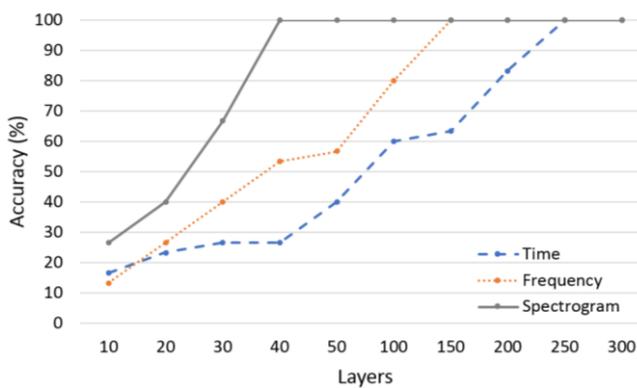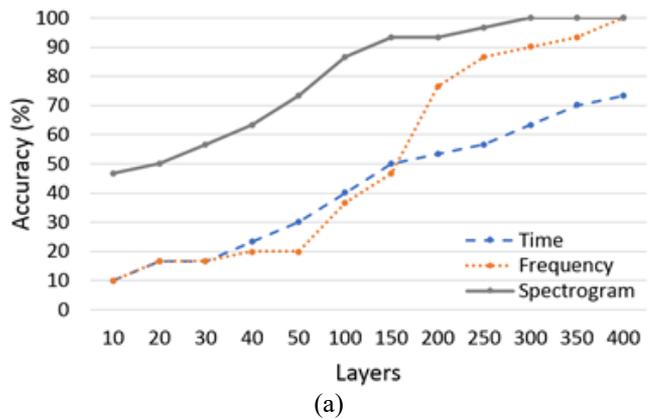| Features Samples	Time Domain	Frequency domain	Spectrogram
10	13.33%	10%	13.33%
20	16.66%	16.66%	20%
30	26.66%	20%	40%
40	26.66%	23.33%	43.33%
50	40%	36.66%	46.66%
100	50%	43.33%	56.66%
150	53.33%	43.33%	80%
200	63.33%	83.33%	83.33%
250	66.66%	90%	96.66%
300	73.33%	93.33%	100%

(a)

(b)

(c)

(d)
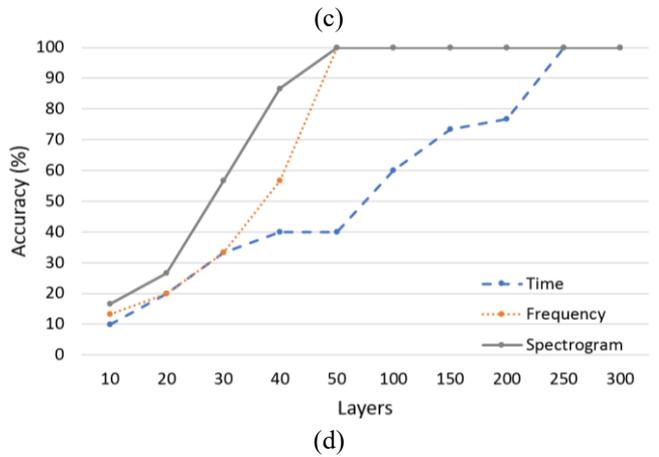
Figure 9. Laboratory test recognition rate for three feature extraction pictures at different layers (a) 15 km/h, (b) 20 km/h, (c) 25 km/h, (d) 30 km/h
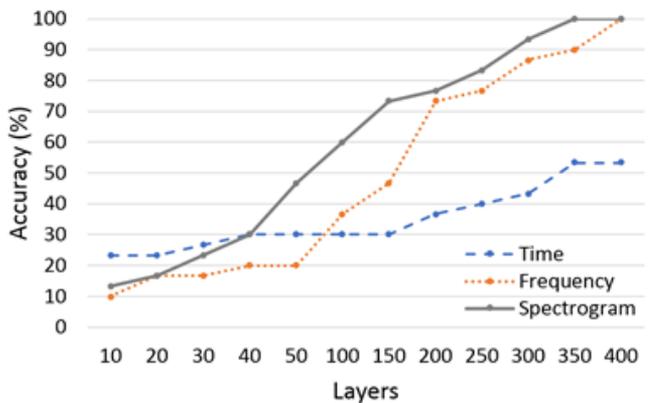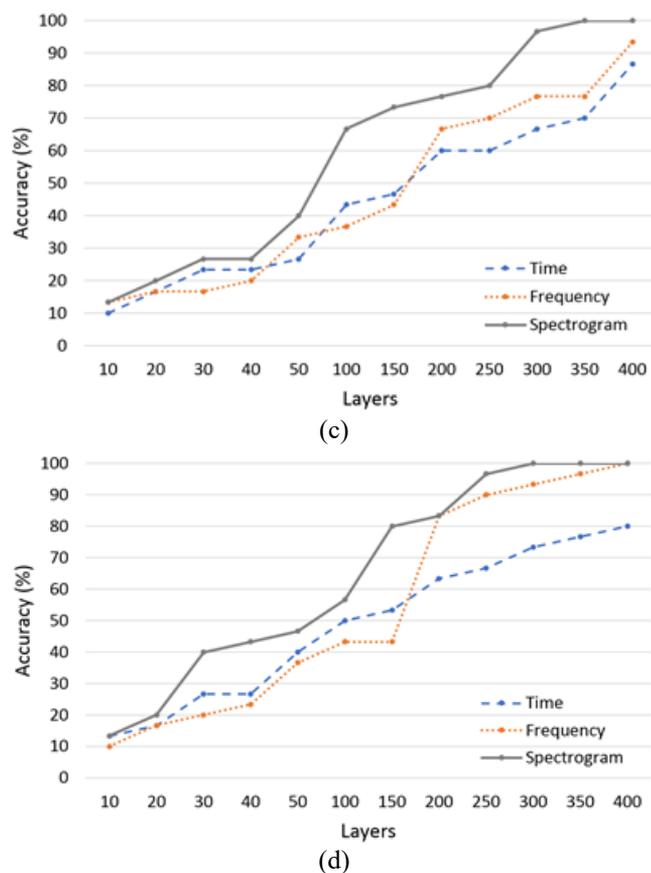

(a)

(b)

(c)

(d)

Figure 10. Road test recognition rate for three feature extraction pictures at different layers (a) 15 km/h, (b) 20 km/h, (c) 25 km/h, (d) 30 km/h

5. CONCLUSION

This study proposed CNN training for different figure recognition to diagnose electric motorbike faults. Four different faults were designed for the electric motorbike, plus a set of sounds under normal conditions. The sound signals of different motorbike speeds during operation were obtained in laboratory and road tests. The analysis included the graph sound signals in the time domain, frequency domain and spectrogram. CNN is used in mechanical learning to train the resulting figures from various states and classify the signal states. Experiments and identification results show that through three different picture presentation methods and CNN, the motorbike fault sound can be effectively identified. The time domain figure without any feature extraction has the lowest recognition rate. The frequency domain figure after the fast Fourier transform has the second highest recognition rate. The spectrogram has the highest recognition rate due to the change in energy. Compared with the actual road test, because the laboratory is less affected by noise, it can have a better identification effect. This problem can be improved by increasing the number of training samples and layers. At the same time, it is proved that this method using the spectrogram figures and CNN can effectively achieve electric motorbike fault diagnosis.

REFERENCES

[1] Ainsworth, W.A. (1976). Automatic speech recognition. Mechanisms of Speech Recognition, 104-119. http://dx.doi.org/10.1016/B978-0-08-020394-2.50014-2

[2] Davis, K.H., Biddulph, R., Balashek, S. (1952). Automatic recognition of spoken digits. The Journal of the Acoustical Society of America, 24(6): 637-642. https://doi.org/10.1121/1.1906946

[3] Denes, P. (1959). The design and operation of the mechanical speech recognizer at University College London. Journal of the British Institution of Radio Engineers, 19(4): 219-229. http://dx.doi.org/10.1049/jbire.1959.0027

[4] Lee, K.F. (1988). Automatic speech recognition: The development of the SPHINX system. 62. Springer Science & Business Media. http://dx.doi.org/10.21437/Interspeech.2018-1085

[5] Minsky, M. (1961). Steps toward artificial intelligence. Proceedings of the IRE, 49(1): 8-30. http://dx.doi.org/10.1109/JRPROC.1961.287775

[6] Newell, A., Simon, H. (1956). The logic theory machine--A complex information processing system. IRE Transactions on Information Theory, 2(3): 61-79. http://dx.doi.org/10.1109/TIT.1956.1056797

[7] Michie, D., Spiegelhalter, D.J., Taylor, C.C. (1994). Machine learning, neural, and statistical classification. Journal of the American Statistical Association, 91(433): 436-38. https://doi.org/10.2307/2291432

[8] Hagan, M.T., Demuth, H.B., Beale, M.H. (1996). Neural Network Design (PWS, Boston, MA). Google Scholar Digital Library. http://dx.doi.org/10.1016/0893-6080(88)90183-9

[9] Zhang, L., Zhou, W., Jiao, L. (2004). Wavelet support vector machine. IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics), 34(1): 34-39. https://dx.doi.org/10.1109/TSMCB.2003.811113

[10] Hinton, G.E., Osindero, S., Teh, Y.W. (2006). A fast learning algorithm for deep belief nets. Neural Computation, 18(7): 1527-1554. http://dx.doi.org/10.1162/neco.2006.18.7.1527

[11] LeCun, Y., Bengio, Y., Hinton, G. (2015). Deep learning. Nature, 521(7553): 436-444. http://dx.doi.org/10.1038/nature14539

[12] LeCun, Y., Boser, B., Denker, J.S., Henderson, D., Howard, R.E., Hubbard, W., Jackel, L.D. (1989). Backpropagation applied to handwritten zip code recognition. Neural Computation, 1(4): 541-551. https://dx.doi.org/10.1162/neco.1989.1.4.541

[13] Janssens, O., Slavkovikj, V., Vervisch, B., Stockman, K., Loccufier, M., Verstockt, S., Van De Walle, R. Van Hoecke, S. (2016). Convolutional neural network based fault detection for rotating machinery. Journal of Sound and Vibration, 377: 331-345. https://dx.doi.org/10.1016/j.jsv.2016.05.027

[14] Li, X., Li, S., Liu, S., Liu, L., He, D. (2020). A malicious webpage detection algorithm based on image semantics. Traitement du Signal, 37(1): 113-118. https://doi.org/10.18280/ts.370115

[15] Fenjiro, Y., Benbrahim, H. (2019) Optimal combination

of imitation and reinforcement learning for self-driving cars. Revue d'Intelligence Artificielle, 33(4): 265-273. https://doi.org/10.18280/ria.330402

[16] Pei, J., Shan, P. (2019). A micro-expression recognition algorithm for students in classroom learning based on convolutional neural network. Traitement du Signal, 36(6): 557-563. https://doi.org/10.18280/ts.360611

[17] Sang, J., Wu, W., Luo, H., Xiang, H., Zhang, Q., Hu, H., Xia, X. (2019). Improved crowd counting method based on scale-adaptive convolutional neural network. IEEE Access, 7: 24411-24419. https://doi.org/10.1109/ACCESS.2019.2899939

[18] Wen, L., Li, X., Gao, L., Zhang, Y. (2017). A new convolutional neural network-based data-driven fault diagnosis method. IEEE Transactions on Industrial Electronics, 65(7): 5990-5998. https://dx.doi.org/10.1109/TIE.2017.2774777

[19] Monteiro, R., Bastos-Filho, C., Cerrada, M., Cabrera, D., Sánchez, R.V. (2018). Convolutional neural networks using fourier transform spectrogram to classify the severity of gear tooth breakage. 2018 International Conference on Sensing, Diagnostics, Prognostics, and Control, pp. 490-496. http://dx.doi.org/10.1109/SDPC.2018.8664985

[20] Bojarski, M., Del Testa, D., Dworakowski, D., Firner, B., Flepp, B., Goyal, P., Jackel L.D., Monfort M., Muller U., Zhang X., Zhao. J., Zieba K. (2016). End to end learning for self-driving cars. https://arxiv.org/abs/1604.07316.

[21] Larminie, J., Lowry, J. (2012). Electric Vehicle Technology Explained. John Wiley & Sons. http://dx.doi.org/10.1002/0470090707.ch7

[22] LeCun, Y., Boser, B.E., Denker, J.S., Henderson, D., Howard, R.E., Hubbard, W.E., Jackel, L.D. (1990). Handwritten digit recognition with a back-propagation network. In Advances in Neural Information Processing Systems, pp. 396-404. http://dx.doi.org/10.1016/S0925-2312(02)00614-8

[23] LeCun, Y., Bottou, L., Bengio, Y., Haffner, P. (1998). Gradient-based learning applied to document recognition. Proceedings of the IEEE, 86(11): 2278-2324. https://dx.doi.org/10.1109/5.726791

[24] Krizhevsky, A., Sutskever, I., Hinton, G.E. (2012). ImageNet classification with deep convolutional neural networks. Communications of the ACM, 60(6): 84-90. http://dx.doi.org/10.1145/3065386

[25] Hoang, D.T., Kang, H.J. (2019). Rolling element bearing fault diagnosis using convolutional neural network and vibration image. Cognitive Systems Research, 53: 42-50. https://dx.doi.org/10.1016/j.cogsys.2018.03.002