

Received October 3, 2020, accepted October 26, 2020, date of current version November 12, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3035157

An Automatic Intelligent Diagnostic Mechanism for the Milling Cutter Wear

BO-LIN JIAN¹, KUAN-TING YU¹, XIAO-YI SU², AND HER-TERNG YAU^{3,4}, (Senior Member, IEEE)

¹Department of Electrical Engineering, National Chin-Yi University of Technology, Taichung 41170, Taiwan

²Graduate Institute of Automation and Technology, National Taipei University of Technology, Taipei 10608, Taiwan

³Department of Mechanical Engineering, National Chung Cheng University, Chiayi 621301, Taiwan

⁴Advanced Institute of Manufacturing with High-tech Innovations, National Chung Cheng University, Chiayi 621301, Taiwan

Corresponding author: Her-Terng Yau (pan1012@ms52.hinet.net)

This work was supported by the Ministry of Science and Technology, Republic of China, Taiwan, under Contract MOST 107-2218-E-167-004, Contract 108-2218-E-194 -013, Contract 109-2221-E-194 -050, and Contract 109-2221-E-167 -010.

ABSTRACT The abrasion of milling cutters is an important factor that affects the accuracy of a workpiece. The intervals between cutter changes is based on the burr condition of the edges on the finished products as well as their dimensional precision. Delayed replacement of cutters will result in a degradation of workpiece quality and it is important that the wear of cutters be monitored in a timely manner. In this study the actual vibration signals generated in a milling process were measured using an Automatic Intelligent Diagnosis Mechanism (AIDM) to determine cutter wear. The AIDM included two feature extraction approaches and three classification methods. The first approach used the Finite Impulse Response Filter (FIR) with Approximate Entropy (ApEn) for feature extraction. The second approach was nonlinear feature mapping using a fractional order Chen-Lee chaotic system. This used chaotic dynamic error centroids and chaotic dynamic error mapping for status identification. After feature extraction the results were substituted into a Back Propagation Neural Network (BPNN), Support Vector Machine (SVM), and a Convolutional Neural Network (CNN) for identification. The results of the experiments showed that a Chaotic Dynamic Error Map of the fractional order Chen-Lee chaotic system in the AIDM had an identification rate of 96.33% using a convolutional neural network. In addition, it was shown that the AIDM model could automatically select the most suitable feature extraction and classification model from the input signal and could determine the wear level milling cutters.

INDEX TERMS Cutter wear, approximate entropy, automatic intelligent diagnosis mechanism, finite impulse response filter, Chen-Lee chaotic system.

I. INTRODUCTION

Tool wear is an unavoidable problem in machine tool use. Damaged or worn tools will result in defects in the finished products and will seriously effect workpiece quality. Excessive tool wear requires extra work to solve surface roughness issues and defects revealed by additional quality inspection of workpieces [1], [2]. Timely inspection and analysis of tool wear is essential to reduce the cost of manpower, cutting tools and workpieces. Tool wear is more clearly defined using ISO stipulated standards to determine the status of cutting tools. The methods for the diagnosis and lifespan analysis of tool wear are all practical. Different types of sensor, such as

accelerometers [3], microphones [4], dynamometers [5]–[7], and others are used to gather the data in actual machining experiments. Some studies have used audio signals generated by the cutting tool during machining. Variations in the vibrations generated by interaction of the tool with the workpiece during machining are classified using a single signal preprocessing method. However, using this method it is difficult to detect over fitting, should this occur, when the amount of data submitted to the trained model for testing has become very large. In such a case identification efficiency of the classification model is reduced. Most of the signal preprocessing methods presently used for tool wear detection transform the time domain signals into frequency domain signals for analysis. For example, when Fourier Transforms [8], [9] and Wavelet Transforms [10], [11] are used for feature extraction,

The associate editor coordinating the review of this manuscript and approving it for publication was Longzhi Yang¹.

the feature mapping capabilities of Fourier Transforms for non-stationary and time-varying signals are limited, and an accurate description of the features of local signals is not possible. Although Wavelet Transform can accurately describe local information, real time processing is difficult to achieve because of the huge amount of computation involved [12].

Therefore, in this study an Automatic Intelligent Diagnostic Mechanism (AIDM), is proposed that captures the vibration signals generated during interaction between a milling cutter and workpiece using a single accelerometer. Feature extraction is done using two different approaches. The first involves a Finite Impulse Response Filter (FIR) [13], [14] for signal decomposition after which feature extraction is done using Approximate Entropy (ApEn) [15]–[17]. The second approach uses a fractional-order Chen-Lee Chaotic system [18], [19] to conduct nonlinear feature mapping and the Chaotic Dynamic Error Centroid and Chaotic Dynamic Error Maps are selected as features of status identification. Finally, the feature extraction data obtained by these two different methods are used for identification by (1) a Back Propagation Neural Network (BPNN) [20], [21], (2) a Support Vector Machine (SVM) [22], [23] and (3) a Convolutional Neural Network (CNN) [24], [25] respectively, to find the most suitable classification model and feature extraction method for signal testing.

II. METHODS

A. EQUIPMENT, EXPERIMENTAL AND RESEARCH STRUCTURE

In this study a QUASER MV154-C CNC milling machine was used as the experimental platform, see FIGURE 1 (a). The device used for measuring the vibration signals generated by cutting action was a Metra Mess-und Frequenztechnik KS943B.100 3-axis accelerometer, see FIGURE 1 (b). The signal acquisition module used was an NI-9234, see FIGURE 1 (c). Specifications of the CNC machine: the maximum stroke of the X, Y, and Z-axes are 762 mm, 530 mm, 560 mm. The frequency range of the KS943B.100 accelerometer is 0.5 Hz to 22,000 H. The signal acquisition module has 4 channels and the sampling rate is 51.2kS/s. The sampling frequency used in the study was 10,000 Hz. The end milling cutter used in this article is a WEENIX model with a 4-flute number of teeth and a diameter of 10mm. The specific processing conditions in this article are: Spindle speed: 3000rpm Feed rate(mm/min): 300 Cutting depth: 0.5mm. The computer used for analysis and identification uses an Intel^(R) Core^(TM) i7-7700 CPU 3.60 GHz, has 16 GB of RAM, and an NVIDIA GeForce GTX-1080 graphics card. Software used was Matlab 2019a with the Machine Learning Toolbox 11.5, the Deep Learning Toolbox 12.1 and the Neural Network Training Toolbox.

The flowchart of the experiments is as shown in FIGURE 2. The process involves the capture of cutter vibration signals from the 3-axis accelerometer during workpiece machining by the signal acquisition module NI-9234. The signal data is processed and identified using the Automatic Intelligent



FIGURE 1. The equipment used in the experiments.

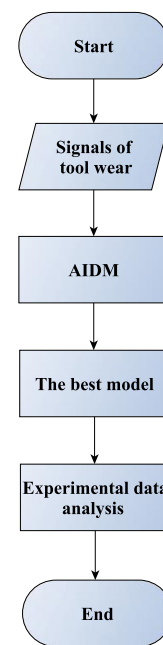


FIGURE 2. Flowchart of the experiment.

Diagnosis Mechanism (AIDM) which will be introduced in section B of part II. The AIDM automatically selects the different extraction methods for the various classification models and applies them to the input data. In this way the most suitable feature extraction method is selected that also gives the best final identification rate and the most suitable classification model. The AIDM used in this study utilizes two different approaches to feature extraction.

The first approach involves the use of a FIR with ApEn, and the fractional order Chen-Lee chaotic system. A transformation between the time domain and the frequency domain

for the vibration wear signals (as slight, moderate, and severe wear) for an end milling cutter is done using the Fast Fourier Transform (FFT). The features are identified as bands that indicate wear status and Bandpass filtering is done for each condition which can be displayed as a frequency domain response diagram. The filtering range is from 2280 Hz to 4500 Hz and the feature bands of the three types of tool wear status can be retained. The regularity of signal feature values of the three states obtained after Bandpass filtering can be distinguished by ApEn. In the second approach nonlinear feature mapping of the time domain vibration signals of each of the three cutter wear states is done using the chaotic system. The actual test signals are then subtracted from the rated ideal signal to obtain a Chaotic Dynamic Error Map with the coordinates of the Chaotic Dynamic Error Centroid as the identification features. In the analyses and comparisons done in the experiments, the y coordinate of the chaotic centroid and the chaotic dynamic error map are used as features for the identification of cutter wear, and thus the number of features is same as in the first approach. Finally, training and identification is done using the three classification models such as BPNN, SVM and CNN. The classification categories include light, moderate, and severe wear.

TABLE 1. Parameters of the training and testing sets.

Experimental set	Training set	Testing set
First set of signals	End mill A	End mill B
Second set of signals	End mills A and B	End mill C

The experimental signals include two sets: the training set and the testing set, as shown in TABLE 1. The first set of signals are the pre-collected milling signals of end mill A and the collected milling signals of end mill B in real time. Feature extraction was conducted for the pre-collected and real-time collected milling signals, and the pre-collected milling signal sets were used for model training and verification. The milling signals collected in real time were used as the testing set. To verify the AIDM performance proposed in this article, the second set of signals was processed in the same way as first set. The second set included the pre-collected milling signals of end mills A and B, and the collected signals from end mill C in real time. The pre-collected milling signals were used as the model training set and for verification. The milling signals collected in real time were used for model testing. According to ISO 8688-2 specifications [26], a milling cutter with an abraded area VB greater than 0.6mm is regarded as damaged. Therefore, in this study a maximum abrasion area of 0.6 mm on an end mill was regarded as severe wear, 0.3 mm as moderate wear, and 0.1 mm as slight wear. The tool wear experiments were carried out based on this classification and the results of the two proposed AIDM feature extraction methods were analyzed and are presented in Sections A and B of part IV. Identification with these two different feature extraction methods was conducted using BPNN, SVM and CNN, and the results of classification and discussions are presented in Sections C and D of part IV.

B. THE AUTOMATIC INTELLIGENT DIAGNOSIS MECHANISM

The AIDM data sources were classified into two categories, the pre-collected signal data and the real-time signal data from the machine. The pre-collected data was used for training and the verification model. The real-time data acquired from the machine was used for status evaluation. The AIDM data processing structure is shown in FIGURE 3. To reduce the classification model training time, the pre-collected data was processed for feature extraction. A portion of the signal features obtained by feature extraction was used for model training and the remainder was used for model verification. After completion of model training, real-time signals from the machine were extracted and used to test the training model. The classification model was compared and ranked according to the test results. The design of the AIDM algorithm used in this study is shown in the flowchart in FIGURE 4.

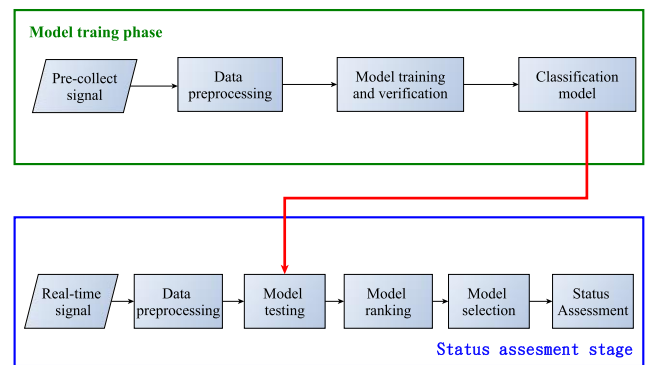


FIGURE 3. The AIDM data processing structure.

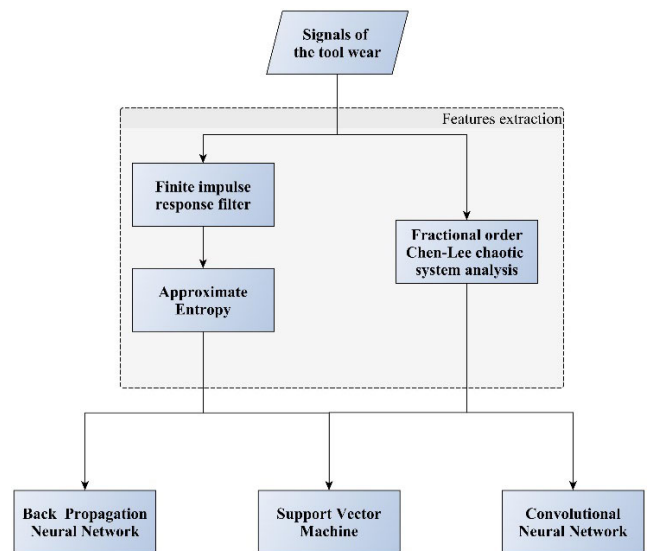


FIGURE 4. The AIDM data processing structure.

III. FEATURE EXTRACTION IN AIDM

A. THE FINITE IMPULSE RESPONSE FILTER

To find the frequency bands features associated with slight, moderate, and severe wear of the tool, transformation of the

vibration signals in the time domain was conducted for the three states using a Fast Fourier Transform (FFT). Bandpass filtering was then done using Finite Impulse Response (FIR). The filtering range used was from 2280 Hz to 4500 Hz to retain the wear features of the frequency bands associated with the three states. The finite impulse response filter, also known as an FIR digital filter, is used to change the frequency spectrum of input signal through computation. The FIR digital filter has a certain distinct advantage in that all the vertexes are within the unit circle after Z conversion. This makes it more stable than the Infinite Impulse Response Filter (IIR) [27]. Furthermore, it does not require feedback in FIR and hence optimization is easier than that for IIR. The input-output function of FIR (1) is as follows:

$$y(n) = \sum_{k=0}^K b_k x(n-k) \tag{1}$$

The input signal is $x(n)$, the output signal of the FIR filter is $y(n)$, K is the filter order, and b_k is the impulse response of the filter which also represents the filter coefficient.

B. APPROXIMATE ENTROPY

In this study Approximate Entropy (ApEn) was used to distinguish the signal difference and regularity of the three states after filtering, and to calculate the signal feature values of the three states. ApEn is a feature extraction method and complexity indicator that can be used to find similarity between segments of signal data as well as data features in numerical form as the difference between data points [28], [29]:

Let $x(1), x(2) \dots x(N)$ represent the data of a time series; where N represents the total number of data points, and the data is formed into an m dimensional vector set $X(i)$ in order, as shown in Formula (2):

$$X(i) = [x(i), x(i+1), \dots, x(i+m-1)], \tag{2}$$

$i = 1, 2, 3 \dots N - m + 1$

$d[X(i), X(j)]$ is the distance between $X(i)$ and $X(j)$, and the maximum distance between these two elements is defined in Formula (3):

$$d[X(i), X(j)] = \max_{k=0 \rightarrow m-1} [|x(i+k) - x(j+k)|] \tag{3}$$

where $1 \leq i \leq N - m + 1$ and $1 \leq j \leq N - m + 1$.

The noise filter coefficient r is given and the number of data that meet the condition of $d[x(i), x(j)] \leq r$ is recorded. The sum of this value is then compared with $N - m + 1$ to obtain the ratio of similar numbers to the total number, as defined in Formula (4):

$$C_i^m(r) = \frac{1}{N - m + 1} \sum_{i=1, i \neq j}^{N-m+1} |d[x(i) - x(j)]| \tag{4}$$

where $d(i, j) = \begin{cases} 1, & |d[x(i) - x(j)]| \leq r \\ 0, & |d[x(i) - x(j)]| \geq r \end{cases}$

The logarithm of $C_i^m(r)$ is used for calculating the average of all i , and it is recorded as $\phi^m(r)$, as defined in Formula (5):

$$\phi^m(r) = \frac{1}{N - m + 1} \sum_{i=1}^{N-m+1} \ln C_i^m(r) \tag{5}$$

Add 1 to the dimension, giving $m + 1$, to obtain $C_i^{m+1}(r)$ and $\phi_i^{m+1}(r)$, and the value of ApEn can be obtained from Formula (6):

$$\text{ApEn}(m, r) = \lim_{N \rightarrow \infty} [\phi^m(r) - \phi^{m+1}(r)] \tag{6}$$

Because the length, N of the actual data point cannot be infinite, the above formula is rewritten as:

$$\text{ApEn}(m, r, N) = \phi^m(r) - \phi^{m+1}(r) \tag{7}$$

C. DYNAMIC ERROR OF THE FRACTIONAL ORDER THE CHEN-LEE CHAOTIC SYSTEM

Chaos theory is part of nonlinear system theory and one of its characteristics is high sensitivity to variations of input signal. The output signal will show significant changes for even the slightest change in input signal. Chaotic systems have strange attractors and their output motion has an orderly and non-periodic trajectory. In this study the Grünwald-Letnikov (G-L) differentiation model was used to enhance the dynamic error variation rate of the chaotic system. The Chen-Lee chaotic system and the Grünwald-Letnikov (G-L) fractional order model are both discussed in this section. The Chen-Lee chaotic system is a chaotic system based on dynamic systems and it is used for non-linear conversion mapping of extracted signals. It uses a Chaotic Dynamic Error Map and the coordinates of Chaotic Dynamic Error Centroid as identification features. The Chen-Lee dynamic equation is as shown in (8).

$$\begin{cases} \dot{x} = -yz + \alpha x \\ \dot{y} = xz + \beta y \\ \dot{z} = \frac{1}{3}xy + \gamma z \end{cases} \tag{8}$$

The above formula is rewritten in the form of the rated ideal signal and the actual testing signal systems, to obtain the dynamic error of the Chen-Lee chaotic system, as shown in (9) and (10).

$$\begin{cases} \dot{x}_m = -yz_m + \alpha x_m \\ \dot{y}_m = xz_m + \beta y_m \\ \dot{z}_m = \frac{1}{3}xy_m + \gamma z_m \end{cases} \tag{9}$$

$$\begin{cases} \dot{x}_s = -yz_s + \alpha x_s \\ \dot{y}_s = xz_s + \beta y_s \\ \dot{z}_s = \frac{1}{3}xy_s + \gamma z_s \end{cases} \tag{10}$$

The dynamic error of the system can be obtained from subtracting the rated ideal signal system from the testing signal system. The above two formulas are rewritten in matrix

form [30], [31] to facilitate subtraction of the two systems, as shown in (11) and (12).

$$\begin{bmatrix} \dot{x}_m \\ \dot{y}_m \\ \dot{z}_m \end{bmatrix} = \begin{bmatrix} \alpha & 0 & -y_m \\ 0 & \beta & x_m \\ 0 & \frac{1}{3}x_m & \gamma \end{bmatrix} \begin{bmatrix} x_m \\ y_m \\ z_m \end{bmatrix} \quad (11)$$

$$\begin{bmatrix} \dot{x}_s \\ \dot{y}_s \\ \dot{z}_s \end{bmatrix} = \begin{bmatrix} \alpha & 0 & -y_m \\ 0 & \beta & x_m \\ 0 & \frac{1}{3}x_s & \gamma \end{bmatrix} \begin{bmatrix} x_s \\ y_s \\ z_s \end{bmatrix} \quad (12)$$

In this article the concept of processing the discrete signal and expressing the time domain vibration signals (Chaotic dynamic) of the rated signal system and the testing signal system is introduced:

$$n[i] = \{n[1], n[2], n[3] \cdots n[i]\} \quad (13)$$

$$t[i] = \{t[1], t[2], t[3] \cdots t[i]\} \quad (14)$$

On the basis of the above mentioned concepts the parameters of the rated signal system as in (11) can be defined as $x_m[i] = n[i + 1]$, $y_m[i] = n[i + 2]$, $z_m[i] = n[i + 3]$ and the parameters of the testing signal system in (12) as $x_s[i] = t[i + 1]$, $y_s[i] = t[i + 2]$, $z_s[i] = t[i + 3]$. Subtract (11) from (12) to obtain the dynamic error of the chaotic system. The dynamic error equation of the chaotic system can be expressed as in (15).

$$\begin{bmatrix} \dot{e}_1[i] \\ \dot{e}_2[i] \\ \dot{e}_3[i] \end{bmatrix} = \begin{bmatrix} \alpha & 0 & 0 \\ 0 & \beta & 0 \\ 0 & 0 & \gamma \end{bmatrix} \begin{bmatrix} e_1 \\ e_2 \\ e_3 \end{bmatrix} + \begin{bmatrix} -e_2e_3 \\ e_1e_3 \\ \frac{1}{3}e_1e_2 \end{bmatrix} \quad (15)$$

The approximate fractional differential equation for the dynamic error of the fractional order chaotic system can be expressed as in Formula (16). Where σ is the order, m is a real number, and the differential computation can be performed when $0 < \sigma < 1$. The dynamic error equation of the fractional order Chen-Lee chaotic system can be obtained from fractional differentiation of Formula (16) and Formula (15) [18], [32], as shown in Formula (17).

$$D_e^{\pm\sigma} e^m \approx \frac{\Gamma(m + 1)}{\Gamma(m + 1 \mp \sigma)} e^{m \mp \sigma} \quad (16)$$

$$\begin{bmatrix} \dot{e}_1[i] \\ \dot{e}_2[i] \\ \dot{e}_3[i] \end{bmatrix} \approx \begin{bmatrix} \alpha' & 0 & 0 \\ 0 & \beta' & 0 \\ 0 & 0 & \gamma' \end{bmatrix} \begin{bmatrix} e_1^{1-\sigma} \\ e_2^{1-\sigma} \\ e_3^{1-\sigma} \end{bmatrix} + \begin{bmatrix} \frac{-\Gamma(1)\dot{e}_2[i]\dot{e}_3[i]\dot{e}_1[i]^{-\sigma}}{\Gamma(1-\sigma)} \\ \frac{\Gamma(1)\dot{e}_1[i]\dot{e}_3[i]\dot{e}_2[i]^{-\sigma}}{\Gamma(1-\sigma)} \\ \frac{\Gamma(1)\dot{e}_1[i]\dot{e}_2[i]\dot{e}_3[i]^{-\sigma}}{3\Gamma(1-\sigma)} \end{bmatrix} \quad (17)$$

$$\begin{cases} \alpha' = \frac{\alpha\Gamma(2)}{\Gamma(2-\sigma)} \\ \beta' = \frac{\beta\Gamma(2)}{\Gamma(2-\sigma)} \\ \gamma' = \frac{\gamma\Gamma(2)}{\Gamma(2-\sigma)} \end{cases} \quad (18)$$

$\dot{e}_1[i] = \dot{x}_m[i] - \dot{x}_s[i]$, $\dot{e}_2[i] = \dot{y}_m[i] - \dot{y}_s[i]$, and $\dot{e}_3[i] = \dot{z}_m[i] - \dot{z}_s[i]$ in Formula (17) and they are defined according to fractional calculus, α' , β' , and γ' are the system parameters as in Formula (18); and $0 < \alpha' < (-\beta' - \gamma')$ is required for Formula (18) to realize the dynamic error of the fractional order Chen-Lee chaotic system [32]. In this article the experimental results are compared and the y coordinates of the chaotic centroid and the chaotic dynamic error map are taken as the features for the analysis of milling cutter wear status.

IV. THE AIDM CLASSIFIER

A. BACK PROPAGATION NEURAL NETWORK

The Back Propagation Neural Network is a combination of Multilayer Perceptron (MLP) [33] and Back Propagation (BP) [34]. The Back Propagation Neural Network is a motoring type of learning network and is the most representative and widely applied neural network. The BPNN includes the forward propagation of network input and back propagation of output errors. If the actual output does not reach the expected level, an error message is sent back from the output layer to the input layer. The purpose of learning can be achieved by successively modifying the weights to reduce the error level.

The basic structure of the Back Propagation Neural Network is shown in FIGURE 5. The structure of the neural network includes the processing unit, the layers, and the network. The processing unit is the most basic computing unit. The “layers” are formed by several processing units with the same function and include the input, hidden and output layer which together form the neural network.

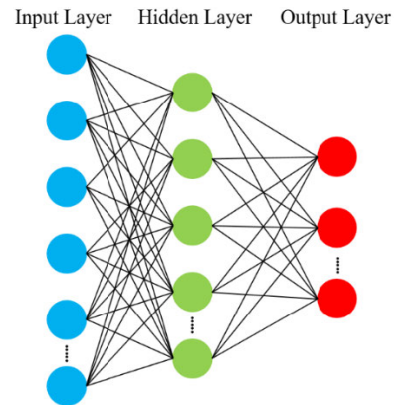


FIGURE 5. Structure of the back propagation neural network.

The Back Propagation Neural Network is a typical multi-layer network structure formed by the input, hidden and output layers. The layers are connected through the computation of weights and errors, and the output value of the previous layer is processed as the input to the next layer with a nonlinear activation function. The formula is as follows:

$$S_j = f \left(\sum_{i=0}^{m-1} w_{ij}x_i + b_j \right) \quad (19)$$

where S_j is the j th input value of the hidden layer unit; w_{ij} is the weight for connecting the processing units in different layers; x_i is the input value of the i th unit; and b_j is a bias value.

When an error is generated between the predicted value and the actual value, the back propagation neuron will pass the error back to the hidden layer to correct the weight and bias value of the model. The input value of the training data will be passed to the next layer based on the new value of weight and bias of each layer, to correct errors until the value is within the acceptable range when it will stop. The objective function to reduce the difference between the output and target value of the neural network, is shown in (20):

$$E = \frac{1}{2} \sum_{i \in \text{outputs}} (y_{(i)} - \hat{y}_{(i)})^2 \quad (20)$$

where $y_{(i)}$ is the actual value; $\hat{y}_{(i)}$ is the predicted value.

B. SUPPORT VECTOR MACHINE

The support vector machine is a widely applied classification technology for machine learning and pattern identification. The support vector machine was first used to solve the two classification issue. The learning approach of the Support Vector Machine learning system is the calculation of an ideal hyperplane through statistical learning theory, so that the data between various categories has the Maximal Margin and the Minimal Misclassification, and the optimal plane consists of the support vectors. H is the classification hyperplane; H_1 and H_2 are the samples that are closest to the classification hyperplane in each of the categories respectively, they are parallel to the classification hyperplane and the distance between them is the margin. For optimal classification of the hyperplane it is necessary that these two types of sample be properly separated to maximize the classification distance.

Where $H_1 = wx + b = 1$, $H_2 = wx + b = -1$ and x is the input feature vector; w is the normal vector of hyperplane; b is a constant and the interval between two hyperplanes is $2 / \|w\|$. Therefore, to maximize the distance

between the two planes, $\|w\|$ needs to be minimized. The following condition must be fulfilled to make the same type of sample data fall together on the same side of the hyperplane:

$$wx_i + b \begin{cases} > 0 & \text{for } y_i = 1 \\ < 0 & \text{for } y_i = -1 \end{cases} \quad (21)$$

The above formula is combined and rewritten as:

$$y_i (wx_i + b) \geq 1, \quad i = 1, 2, 3 \dots, n \quad (22)$$

To minimize $\|w\|$ and fulfill the conditional expression to achieve the optimal classification hyperplane, we combine the above expressions and obtain the following objective function:

$$\begin{aligned} & \min \frac{1}{2} \|w\|^2 \\ & \text{subject to } \{y_i [(wx_i + b)] - 1 \geq 0, \quad (i = 1, 2, 3 \dots, n) \end{aligned} \quad (23)$$

C. CONVOLUTIONAL NEURAL NETWORK

The Convolutional Neural Network (CNN) is a deep learning network that is capable of high-performance image processing. Its characteristic is that the feature of each layer in the convolutional neural network is obtained through convolution and an activation function after full-value sharing by the local area of the previous layer. The CNN reduces complexity by a reduction of neurons and feature extraction which also significantly reduces the training parameters for images within a certain dimension [35], [36]. CNN can use an original image directly as input to the network and this results in its wide application in the field of deep learning. The convolutional neural network has Input, Convolution, Pooling, Activation, Fully Connected and Output Layers [37]. FIGURE 6 is a diagram of the convolutional neural network designed for this study. It was used as a classification model for subsequent AIDM.

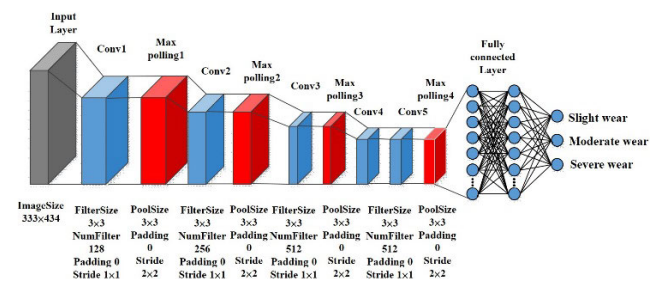


FIGURE 6. Structure of the convolutional neural network.

The convolutional layer extracts image features and conducts convolution computations using masks of different sizes for spatial filtering to enhance image features and extraction. In general, the pooling layer includes the Max Pooling Layer [38] and Average Pooling Layer [39]. The activation layer performs nonlinear conversion through the activation function to obtain the features. Commonly seen activation functions include the Sigmoid (24) and ReLU (25). However, the Sigmoid function has a gradient loss problem. In contrast, the output of ReLU is relatively stable [40]. Therefore, to retain the most significant features of the images, prevent gradient loss and increase the computation speed of the network, maximum pooling and ReLU were used for parameter settings.

$$f(x) = \frac{1}{1 + e^{-x}} \quad (24)$$

$$f(x) = \begin{cases} 0 & \text{for } x < 0 \\ x & \text{for } x \geq 0 \end{cases} \quad (25)$$

V. EXPERIMENTAL RESULTS AND DISCUSSION

A. FREQUENCY SPECTRUM ANALYSIS AND FEATURE EXTRACTION ENTROPY

FIGURE 7 (a-c) shows original vibration signals from milling operations carried out with end mills that had with slight, moderate and severe wear, see the training sets in TABLE 1. The time domain and frequency domain signals

were converted using a FFT and the feature band for the various states were determined according to the frequency domain response diagram, see FIGURE 8 (a-c). It can be seen that the feature band of slight wear was 3272 to 3344 Hz, that of moderate wear was 4292 to 4395 Hz, and that of severe wear was 3105 Hz to 3223 Hz. A Bandpass filter was used to filter all signals between 2280 and 4500 Hz to facilitate subsequent feature extraction and classifier identification. Feature extraction was done after Bandpass filtration, the ratio of similar numbers to the total number was calculated and dimensionality was reduced all using ApEn computation.

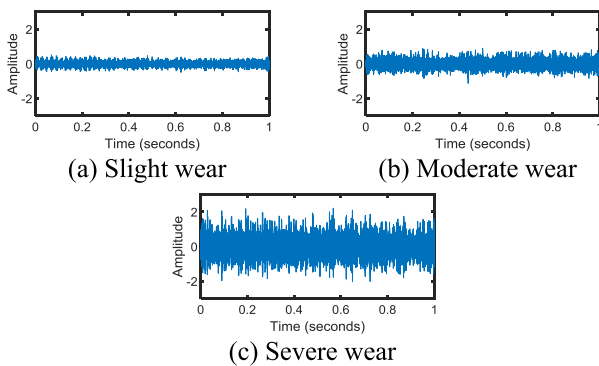


FIGURE 7. Time domain signals from cutters in three different states of wear.

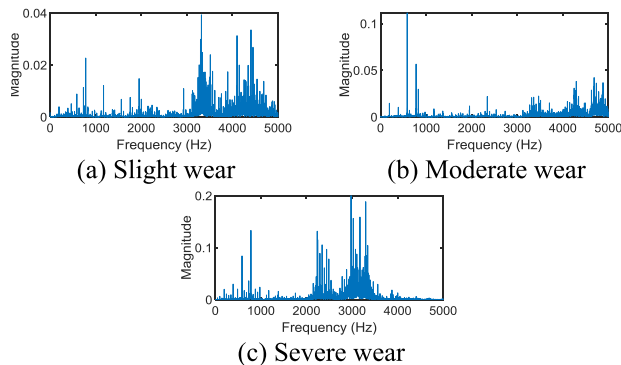


FIGURE 8. Frequency domain signal diagram of the three wear states.

FIGURE 9 is the feature signal map after feature extraction with ApEn. It can be seen that although the difference between slight wear and moderate wear with FIR and ApEn can be seen, the curve traces of moderate and severe wear are very similar, these two states cannot be independent, and this may cause classifier error.

B. NONLINEAR FEATURE EXTRACTION OF THE CHEN-LEE CHAOTIC SYSTEM FRACTIONAL ORDER

Nonlinear feature mapping of the Time Domain signals of the three states of the tool in FIGURE 7 (a-c) was conducted and the sensitivity of the chaotic system allowed the different states to be individually separated by slight changes of signal intensity. The experimental results were compared and the y-axis coordinates of the chaotic centroid and the chaotic

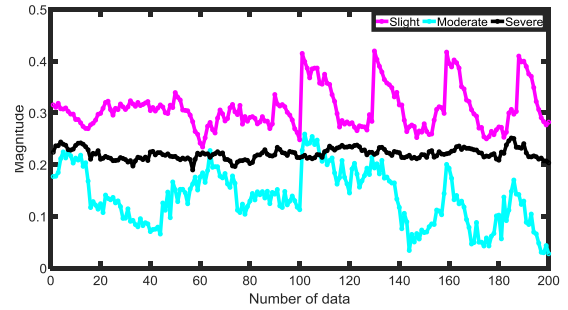


FIGURE 9. Signal distribution of the three feature states for the second signal training set after ApEn computation.

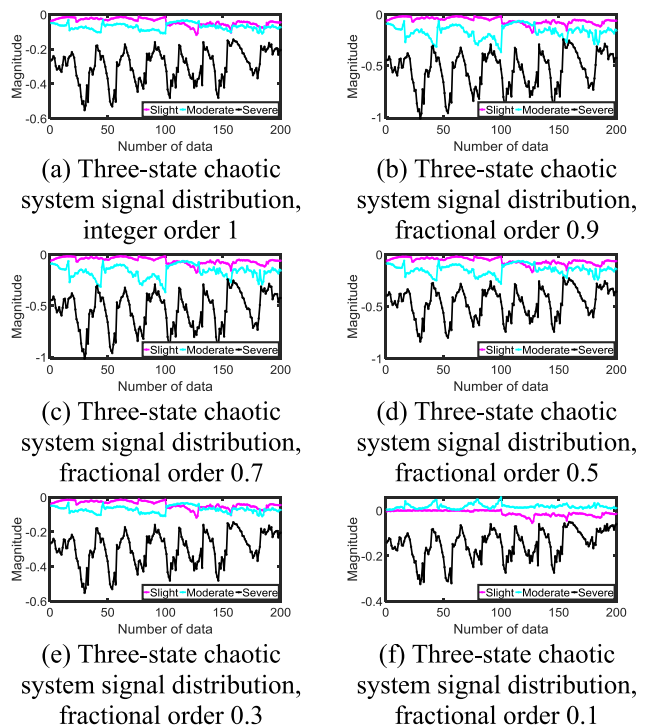


FIGURE 10. Three-state chaotic system signal distribution, integer and fractional orders.

dynamic error map were taken as the physical quantity for the analysis of tool wear. The distribution of the feature signals for integer order and fractional order were also compared, see FIGURE 10. As the feature signal distribution map of the integer order 1 to the fractional order 0.2 showed insignificant change, the feature signal distribution maps using integer order 1, fractional orders 0.9, 0.7, 0.5, 0.3 and 0.1 were all tried.

The feature distribution in the three states were compared, see FIGURE 10. When the order was 0.1, the feature distribution of the three different wear states were independent and the size and distribution interval of each could be clearly identified. This was not the case with the other orders. Therefore, the fractional order 0.1 was used as the parameter for subsequent experiments. The chaotic dynamic error map and centroids were substituted into the classifier for subsequent training and testing. The dynamic error map

of the chaotic system with fractional order 0.1 for the three cutter wear states is shown in FIGURE 11(a-c). The difference between the three wear states after nonlinear mapping through the fractional order Chen-Lee chaotic system can be clearly seen and the wear state of a milling cutter can be easily determined.

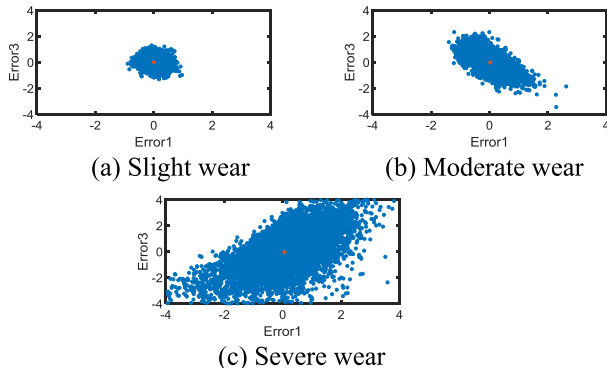


FIGURE 11. Dynamic error maps of the chaotic system with fractional order 0.1.

C. TRAINING AND IDENTIFICATION, BPNN AND SVM

There are no fixed specifications or basis for the internal parameter settings of the Back Propagation Neural Network, so the constant parameters used for training, the number of neurons, and the amount of data training were used. In this study, the training function used in BPNN for AIDM was a Scale Conjugate Gradient Algorithm. The GPU was also used for conducting auxiliary computations and the number of neurons was set to 1000. In TABLE 2 it can be seen that the first set of experimental signals were 300 milling data signals from end mill A for model training; 300 real-time data signals were also collected from end mill B and substituted into the trained model for testing. Using the same parameters, the second set of experimental signals sent 600 pre-collected milling data signals from end mills A and B for model training. The 300 collected real-time data signals from end mill C were used as the testing set and substituted into the trained model for testing.

TABLE 2. Identification of BPNN tool wear for the first and second set of signals.

Experimental signal set	Feature extraction method	Number of training data	Number of testing data	Identification rate (%)
First set	FIR+ApEn	300	300	89.33
First set	Fractional order Chen-Lee Chaotic System	300	300	98.66
Second set	FIR+ApEn	600	300	74.33
Second set	Fractional order Chen-Lee Chaotic System	600	300	85.33

The features obtained from the computation of ApEn and the y-axis coordinates of the chaotic centroids with the

fractional order 0.1 are substituted into BPNN for training and testing using the same parameters and data as in the experiment. The difference in identification rate between the preprocessing methods for the two types of data was compared and the results can be seen in TABLE 2.

The experimental parameters for SVM are same as for the BPNN. The pre- collected milling signals were used as the training set and sent for model training, and the real-time collected signals were used as the testing set. The features obtained from the computation of ApEn and the y-axis coordinates of the chaotic centroids with the fractional order 0.1 were substituted for training and testing just as for the BPNN. The difference in identification rate between these two types of feature extraction methods was compared, see TABLE 3.

TABLE 3. The SVM identification of tool wear for the first and second set of signals.

Experimental signal set	Feature extraction method	Number of training data	Number of testing data	Identification rate (%)
First set	FIR+ApEn	300	300	89
First set	Fractional order Chen-Lee Chaotic System	300	300	97.667
Second set	FIR+ApEn	600	300	79.67
Second set	Fractional order Chen-Lee Chaotic System	600	300	83.33

D. TRAINING AND IDENTIFICATION OF THE CONVOLUTIONAL NEURAL NETWORK

Matlab 2019a was used in this study for verification and the Deep Network Designer was used for the implementation of a self-made convolutional neural network as well as the verification classifier, see FIGURE 6. The dynamic error map of the chaotic system with fractional order 0.1 for the three states of tool was substituted into the convolutional neural network for training and testing. In this section the training and testing sets from the second set of signals was used as an example. TABLE 4 shows the number of samples in the CNN training set. The training set samples were substituted into the convolutional neural network for training, and the classification identification rate and the number of iterations are shown in FIGURE 12.

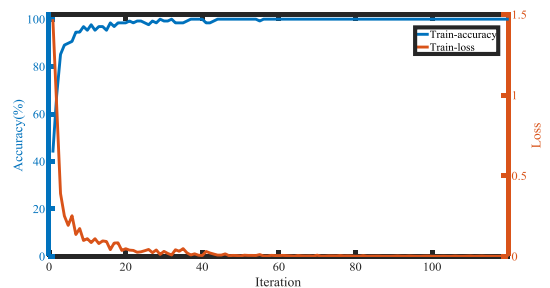


FIGURE 12. The convolutional neural network training and iterative process.

TABLE 4. The number of samples in the CNN training set.

Item	Number of pictures (pieces)
Slight wear	200
Moderate wear	200
Severe wear	200

In Figure 12 it can be seen that the accuracy rate reaches 100% after 60 iterations and the loss function is 0. This indicates that the model has been completely trained. The images to be tested are substituted into the trained model and the number of image samples in the second CNN testing set were as shown in Table 5.

TABLE 5. The number of samples in the CNN testing set.

Item	Number of pictures (pieces)
Slight wear	100
Moderate wear	100
Severe wear	100

The results of substituting the testing set images from the first and second set of signals into the convolutional neural network model are shown in TABLE 6, presented in the form of the number of test samples and the number of correct identifications. The identification rates were 96.33% and 89.33% respectively.

TABLE 6. The corresponding identification results of CNN tool wear for experimental signals.

Experimental signal set	Number of training samples	Number of test samples	Number of correct identifications	Identification rate (%)
First set	300	300	289	96.333
Second set	600	300	268	89.333

VI. CONCLUSION

In this study the AIDM framework was used in an investigation of the wear rate of end milling cutters used in production. Two different signal preprocessing approaches were used to carry out feature extraction and subsequent classifier identification. The first approach used FIR filtration with ApEn. Although this approach could be used to successfully detect differences between the state of slight and moderate wear, the results did not show independency and this could cause errors in classification. The second approach used a fractional order Chen-Lee chaotic system. The characteristic high sensitivity of a chaotic system to small changes of input proved advantageous. Feature mapping was done and the chaotic dynamic error maps and the y-axis coordinates of the chaotic centroids were used as identification features. The features extracted with these two approaches were used in the classifier for subsequent identification. The results showed that the AIDM designed in this study had an identification rate of 98.66% with the BPNN classifier for the first signal test set. For the second test set the identification rate was 89.33% using the CNN classifier. In both approaches the Chen-Lee chaotic system was used for fractional order signal extraction. The fractional order Chen-Lee chaotic system and the AIDM

designed in this article gave the best and most effective feature extraction. The fractional order Chen-Lee chaotic system combined with the convolutional neural network also showed good results with the self-designed AIDM structure. The identification rates achieved with the two signal test sets was more than 89%. The results of the experiment also showed that the AIDM method could provide a most suitable classification model as well as excellent signal preprocessing.

REFERENCES

- [1] B. Jinsong, G. Yuan, Z. Xiaohu, Z. Jianguo, and J. Xia, "A data driven model for predicting tool health condition in high speed milling of titanium plates using real-time SCADA," *Procedia CIRP*, vol. 61, pp. 317–322, Apr. 2017.
- [2] Y. Dai and K. Zhu, "A machine vision system for micro-milling tool condition monitoring," *Precis. Eng.*, vol. 52, pp. 183–191, Apr. 2018.
- [3] B. S. Prasad and M. P. Babu, "Correlation between vibration amplitude and tool wear in turning: Numerical and experimental analysis," *Eng. Sci. Technol., Int. J.*, vol. 20, no. 1, pp. 197–211, Feb. 2017.
- [4] J. Kopač and S. Šali, "Tool wear monitoring during the turning process," *J. Mater. Process. Technol.*, vol. 113, nos. 1–3, pp. 312–316, Jun. 2001.
- [5] R. E. Haber, J. E. Jiménez, C. R. Peres, and J. R. Alique, "An investigation of tool-wear monitoring in a high-speed machining process," *Sens. Actuators A, Phys.*, vol. 116, no. 3, pp. 539–545, Oct. 2004.
- [6] M. Malekian, S. S. Park, and M. B. G. Jun, "Tool wear monitoring of micro-milling operations," *J. Mater. Process. Technol.*, vol. 209, no. 10, pp. 4903–4914, Jun. 2009.
- [7] J. Downey, S. Bombinski, M. Nejman, and K. Jemielniak, "Automatic multiple sensor data acquisition system in a real-time production environment," *Procedia CIRP*, vol. 33, pp. 215–220, Jan. 2015.
- [8] C. S. MacInnes, "FFT-based joint bearing estimation," *IEEE J. Ocean. Eng.*, vol. 27, no. 4, pp. 774–779, Oct. 2002.
- [9] J. D. J. Rangel-Magdaleno, R. D. J. Romero-Troncoso, R. A. Osornio-Rios, E. Cabal-Yepez, and A. Dominguez-Gonzalez, "FPGA-based vibration analyzer for continuous CNC machinery monitoring with fused FFT-DWT signal processing," *IEEE Trans. Instrum. Meas.*, vol. 59, no. 12, pp. 3184–3194, Dec. 2010.
- [10] W. Deng, S. Zhang, H. Zhao, and X. Yang, "A novel fault diagnosis method based on integrating empirical wavelet transform and fuzzy entropy for motor bearing," *IEEE Access*, vol. 6, pp. 35042–35056, 2018.
- [11] Z. Huo, Y. Zhang, P. Francq, L. Shu, and J. Huang, "Incipient fault diagnosis of roller bearing using optimized wavelet transform based multi-speed vibration signatures," *IEEE Access*, vol. 5, pp. 19442–19456, 2017.
- [12] H. Ding, D. Liu, and F.-F. Zhao, "Variation trend analysis of groundwater depth in area of well irrigation in Sanjiang plain based on wavelet neural network," in *Advances in Neural Network Research and Applications*, Z. Zeng and J. Wang, Eds. Berlin, Germany: Springer, 2010, pp. 829–836.
- [13] B. N. Mohapatra and R. K. Mohapatra, "Performance analysis on frequency response of finite impulse response filter," *Procedia Comput. Sci.*, vol. 79, pp. 729–736, Jan. 2016.
- [14] Y. A. Grebenko and R. I. Polyak, "Choosing a FIR filter order for the linearization of an analog LPF phase response," in *Proc. Syst. Signal Synchronization, Generating Process. Telecommun. (SYNCHROINFO)*, Jul. 2018, pp. 1–3.
- [15] X. Gao, X. Yan, P. Gao, X. Gao, and S. Zhang, "Automatic detection of epileptic seizure based on approximate entropy, recurrence quantification analysis and convolutional neural networks," *Artif. Intell. Med.*, vol. 102, Jan. 2020, Art. no. 101711.
- [16] W.-L. Chu, M.-W. Huang, B.-L. Jian, and K.-S. Cheng, "Analysis of EEG entropy during visual evocation of emotion in schizophrenia," *Ann. Gen. Psychiatry*, vol. 16, no. 1, p. 34, Dec. 2017.
- [17] Z. Junyi, Y. Chaoying, X. Zhiwei, L. Yue, and W. Zhiqi, "A method for identifying the fault current of DC traction power supply system based on EMD approximate entropy," in *Proc. Int. Conf. Green Energy Appl. (ICGEA)*, Mar. 2017, pp. 18–21.
- [18] H.-T. Yau, S.-Y. Wu, C.-L. Chen, and Y.-C. Li, "Fractional-order chaotic Self-Synchronization-Based tracking faults diagnosis of ball bearing systems," *IEEE Trans. Ind. Electron.*, vol. 63, no. 6, pp. 3824–3833, Jun. 2016.

- [19] B.-L. Jian, C.-C. Wang, J.-Y. Chang, X.-Y. Su, and H.-T. Yau, "Machine tool chatter identification based on dynamic errors of different self-synchronized chaotic systems of various fractional orders," *IEEE Access*, vol. 7, pp. 67278–67286, 2019.
- [20] A. K. Singh, S. S. Panda, D. Chakraborty, and S. K. Pal, "Predicting drill wear using an artificial neural network," *Int. J. Adv. Manuf. Technol.*, vol. 28, nos. 5–6, pp. 456–462, Mar. 2006.
- [21] S. Garg, S. K. Pal, and D. Chakraborty, "Evaluation of the performance of backpropagation and radial basis function neural networks in predicting the drill flank wear," *Neural Comput. Appl.*, vol. 16, p. 407, Jul. 2006.
- [22] B. Kaya, C. Oysu, H. M. Ertunc, and H. Ocak, "A support vector machine-based online tool condition monitoring for milling using sensor fusion and a genetic algorithm," *Proc. Inst. Mech. Eng., Part B, J. Eng. Manuf.*, vol. 226, no. 11, pp. 1808–1818, Nov. 2012.
- [23] J. Karandikar, "Machine learning classification for tool life modeling using production shop-floor tool wear data," *Procedia Manuf.*, vol. 34, pp. 446–454, Jun. 2019.
- [24] S. Kiranyaz, A. Gastli, L. Ben-Brahim, N. Al-Emadi, and M. Gabbouj, "Real-time fault detection and identification for MMC using 1-D convolutional neural networks," *IEEE Trans. Ind. Electron.*, vol. 66, no. 11, pp. 8760–8771, Nov. 2019.
- [25] C.-H. Lee, J.-S. Jwo, H.-Y. Hsieh, and C.-S. Lin, "An intelligent system for grinding wheel condition monitoring based on machining sound and deep learning," *IEEE Access*, vol. 8, pp. 58279–58289, 2020.
- [26] L. L. Alhadeff, M. B. Marshall, D. T. Curtis, and T. Slatter, "Protocol for tool wear measurement in micro-milling," *Wear*, vols. 420–421, pp. 54–67, Feb. 2019.
- [27] T. Tarumi, G. W. Small, R. J. Combs, and R. T. Kroutil, "Infinite impulse response filters for direct analysis of interferogram data from airborne passive Fourier transform infrared spectrometry," *Vibrational Spectrosc.*, vol. 37, no. 1, pp. 39–52, Jan. 2005.
- [28] L. Li, W. Chen, X. Shao, and Z. Wang, "Analysis of amplitude-integrated EEG in the newborn based on approximate entropy," *IEEE Trans. Biomed. Eng.*, vol. 57, no. 10, pp. 2459–2466, Oct. 2010.
- [29] R. Hornero, D. Alvarez, D. Abasolo, F. D. Campo, and C. Zamarron, "Utility of approximate entropy from overnight pulse oximetry data in the diagnosis of the obstructive sleep apnea syndrome," *IEEE Trans. Biomed. Eng.*, vol. 54, no. 1, pp. 107–113, Jan. 2007.
- [30] Y.-C. Kuo, C.-T. Hsieh, H.-T. Yau, and Y.-C. Li, "Research and development of a chaotic signal synchronization error dynamics-based ball bearing fault diagnoser," *Entropy*, vol. 16, no. 10, pp. 5358–5376, Oct. 2014.
- [31] C.-J. Lin, X.-Y. Su, K.-T. Yu, B.-L. Jian, and H.-T. Yau, "Inspection on ball bearing malfunction by chen-lee chaos system," *IEEE Access*, vol. 8, pp. 28267–28275, 2020.
- [32] H.-T. Yau, C.-C. Wang, J.-Y. Chang, and X.-Y. Su, "A study on the application of synchronized chaotic systems of different fractional orders for cutting tool wear diagnosis and identification," *IEEE Access*, vol. 7, pp. 15903–15911, 2019.
- [33] M. Heidari and H. Shamsi, "Analog programmable neuron and case study on VLSI implementation of multi-layer perceptron (MLP)," *Microelectron. J.*, vol. 84, pp. 36–47, Feb. 2019.
- [34] S.-H. Oh, "Error back-propagation algorithm for classification of imbalanced data," *Neurocomputing*, vol. 74, no. 6, pp. 1058–1061, Feb. 2011.
- [35] Y. Wang, H. Li, and X. Li, "A case of on-chip memory subsystem design for low-power CNN accelerators," *IEEE Trans. Comput.-Aided Design Integr. Circuits Syst.*, vol. 37, no. 10, pp. 1971–1984, Oct. 2018.
- [36] L. Koskinen, A. Paasio, and K. A. I. Halonen, "Motion estimation computational complexity reduction with CNN shape segmentation," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 15, no. 6, pp. 771–777, Jun. 2005.
- [37] J. Wang, J. Lin, and Z. Wang, "Efficient hardware architectures for deep convolutional neural network," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 65, no. 6, pp. 1941–1953, Jun. 2018.
- [38] F. Guo, R. He, and J. Dang, "Implicit discourse relation recognition via a BiLSTM-CNN architecture with dynamic chunk-based max pooling," *IEEE Access*, vol. 7, pp. 169281–169292, 2019.
- [39] S. Wang, Y. Jiang, X. Hou, H. Cheng, and S. Du, "Cerebral micro-bleed detection based on the convolution neural network with rank based average pooling," *IEEE Access*, vol. 5, pp. 16576–16583, 2017.
- [40] G. Wang, G. B. Giannakis, and J. Chen, "Learning ReLU networks on linearly separable data: Algorithm, optimality, and generalization," *IEEE Trans. Signal Process.*, vol. 67, no. 9, pp. 2357–2370, May 2019.



BO-LIN JIAN received the B.S. degree from the Department of Electrical Engineering, National Formosa University, in 2009, the M.S. degree in materials science and engineering from the National Taiwan University of Science and Technology, in 2011, and the Ph.D. degree from the Department of Aeronautics and Astronautics, National Cheng Kung University, in 2017.

He is currently an Assistant Professor with the Department of Electrical Engineering, National Chin-Yi University of Technology, Taichung, Taiwan. His current research interests include signal and image processing, machine learning, and control systems.



KUAN-TING YU received the B.S. and M.S. degrees from the Department of Electrical Engineering, National Chin-Yi University of Technology, in 2018 and 2020, respectively. His research interest includes nonlinear system analysis and control.



XIAO-YI SU received the B.S. and M.S. degrees from the National Chin-Yi University of Technology, Taichung, Taiwan, in 2017, and 2019, respectively. He is currently pursuing the Ph.D. degree with the College of Mechanical and Electrical Engineering, National Taipei University of Technology. His research interest includes nonlinear system analysis and control.



HER-TERNG YAU (Senior Member, IEEE) received the B.S. degree from National Chung Hsing University, Taichung, Taiwan, in 1994, and the M.S. and Ph.D. degrees from National Cheng Kung University, Tainan, Taiwan, in 1996 and 2000, respectively, all in mechanical engineering. He is currently a Professor with the Department of Mechanical Engineering, National Chung Cheng University, Chiayi, Taiwan. He has authored more than 150 research articles on a wide variety of topics in mechanical and electrical engineering. His research interests include energy converter control, the system control of mechatronics, and nonlinear system analysis and control.

...