A CNN-Regression-Based Contact Erosion Measurement Method for AC Contactors

Ziran Wu (吴自然), Chufu Fang (方初富), Guichu Wu (吴桂初), Zhenquan Lin (林振权), and Wei Chen (陈威)

***Abstract*— This article proposes a method that indirectly measures the contact erosion of alternating current (ac) contactors via mapping electrical signals to the contacting alloy erosion condition which is represented by the accumulated contact mass loss (ACML). Electrical signal waveforms and their corresponding ACMLs of every make-and-break operation are acquired in endurance tests. A supervised convolutional neural network regression (CNNR) architecture containing six 1-D convolution layers is proposed to model the relation between waveforms and ACMLs. We compare different CNNR architectures as well as different training schemes by the test precision to obtain the optimal solution. Experiments prove that the proposed CNNR architecture with an optimized training scheme can achieve a precise ACML measurement when only voltage waveforms of make operations are used. The best results reach mean absolute errors of 3.29% and 1.59% corresponding to two datasets, respectively, which are superior to the results of other regression methods in the comparison and prove the theoretical significance and application values.**

***Index Terms*— Alternating current (ac) contactors, contact erosion measurement, convolutional neural network regression (CNNR).**

I. Introduction

ALTERNATING current (ac) contactors, including electromagnetic and permanent-magnetic types, are electrical apparatuses used to switch ON and OFF circuits by control signals [1]. AC contactors that perform switching by contacts are the most widely used type for circuits with large currents. Although the technology of solid-state contactors has been developing quickly in recent years, the applications of solid-state contactors are limited since, at present, their current capacity is still far lower than that of traditional models.

AC contactors are most widely utilized in low-voltage (lower than ac 1200 V) applications, especially for high-power motor control. Hence, this article investigates ac contactors in low-voltage circumstances. In an ac contactor, the driving system drives the contacts to connect (termed "make") and separate (termed "break") for circuit switching ON and OFF. The electrical endurance, termed the electrical lifetime/durability, is determined by the residual contacting alloy (normally $AgSnO_2$ or AgCdO) on their contacts (see Fig. 1 where the structure of an unused contact is shown). The alloy is of low resistance, high hardness, and high thermostability and is used to improve the connective performance and suppress the erosion rate. Unlike circuit breakers [2], switching operations are executed much more frequently with ac contactors and therefore the alloy will be eroded quickly. A contact failure occurs when the alloy is completely eroded, and a failed contact causes welding or lose connectivity that leads to circuit faults. Fig. 2 shows all contacts of an ac contactor after an ac-4 test [1]. In Fig. 2, the photographs in each

Fig. 1. Unused contact.

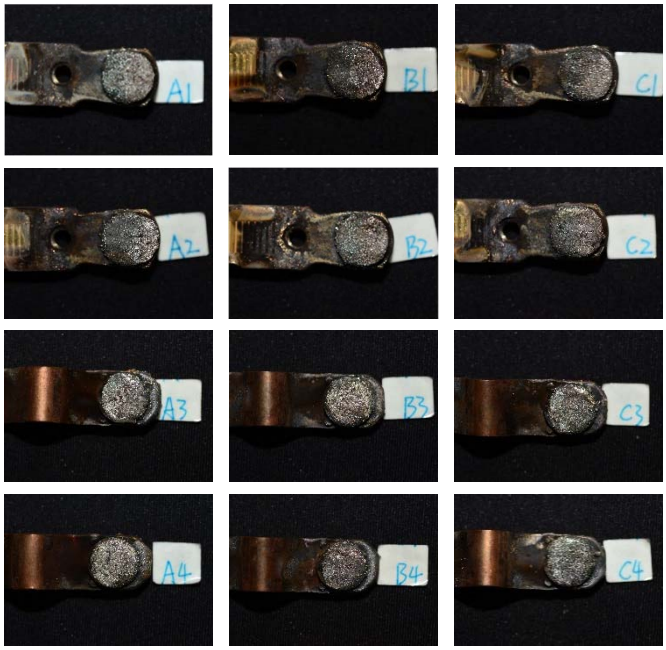

Fig. 2. Contact erosion after an ac-4 test contact carrier, contacting alloy.

Manuscript received 1 April 2022; revised 16 June 2022; accepted 6 July 2022. Date of publication 19 July 2022; date of current version 2 August 2022. This work was supported in part by the National Natural Science Foundation of China under Grant 51507113, in part by the Wenzhou Municipal Science and Technology Bureau under Grant ZG2019017 and Grant ZG2020049, and in part by the Zhejiang Provincial Natural Science Foundation under Grant LGG20E070005. The Associate Editor coordinating the review process was Hongrui Wang. *(Corresponding author: Guichu Wu.)*

Ziran Wu, Zhenquan Lin, and Wei Chen are with the Engineering Research Center of Low-Voltage Apparatus Technology of Zhejiang Province, College of Electrical and Electronic Engineering, Wenzhou University, Wenzhou, Zhejiang 325035, China (e-mail: naturex@wzu.edu.cn; zqlin@wzu.edu.cn; chenwei1989@wzu.edu.cn).

Chufu Fang is with Zhengjiang Tengen Electric Company Ltd., Wenzhou, Zhejiang 325699, China (e-mail: fangchufu@tengen.com.cn).

Guichu Wu is with Zhejiang Juchuang Smartech Company Ltd., Wenzhou, Zhejiang 325036, China (e-mail: wgc@wzu.edu.cn).

Digital Object Identifier 10.1109/TIM.2022.3192282

column show the four contacts of a phase. The erosion is occurred in both making and breaking operations. In the making process, a strong collision as well as several bounces with arc discharges occur between the alloy parts of two contacts. In the breaking process, a strong arc discharge with a high temperature (higher than 6000 K) happens when two alloy parts separate. These phenomena are the main causes of erosion.

Hence, contact erosion measurement is eagerly required, since users expect to monitor the residual electrical endurance so that they can repair or replace ac contactor right before their failure occurrences to guarantee the reliability of frequently switching circuits. To evaluate the erosion condition of the alloy, the most straightforward approach is to directly measure the thickness or the mass of the alloy part. However, it is impractical to frequently disassemble an operating ac contactor for measurement. Hence, an online measurement method not disturbing the working circumstance is expected. The traditional approach widely utilized in the past is to indicate the erosion condition by counting the number of operations, which is currently obsoleted owning to a critical weakness: nonconstant load circumstance leads to varying erosion rate, which cannot be reflected by simply counting operations. Improved approaches normally apply acquired signals from long-term monitoring to achieve erosion estimation. However, using long-term monitoring data leads to the loss of real-time capability.

Hence, this article proposes a novel contacting alloy erosion measurement method employing convolutional neural network regression (CNNR) with a specially designed 1-D CNNR architecture. Compared with the existing methods, the proposed method has outstanding advantages in the real-time measurement capability, as well as in precision, flexibility, and practicality.

II. Prior Work

A. Review of the State of the Art

Erosion condition measurement approaches for different types of electrical apparatuses, for example, contactors, circuit breakers, relays, and so on, usually share common methods and technologies. Some scholars consider the erosion condition measurement as a probability problem and model it by the reliability distribution. Zheng *et al.* [3] illustrated that the manufacturing dispersion degree determines the electrical life distribution whether be the three-parameter Weibull distribution or the normal distribution. Li *et al.* [4] proposed an online prediction method that estimated the contact erosion loss by a Wiener-based degradation model and established the probability density function of the residual electrical lifetime. This type of approach achieves probabilistic electrical endurance under a general statistic pattern. However, the peculiarities of an apparatus, that is, the individual quality, the working circumstance, and so on, are not reflected in the estimation.

Mapping the contact condition of an electrical apparatus from signals acquired during its operating time is another approach that can distinguish different individuals and be practically applied in industry. In other words, users can achieve remote, online erosion measurement by installing communicable sensors and processing acquired signals by computers. Considering that the characteristics are not significant when the contacts are solidly connected or completely separated, researchers focus on the fluctuation signals generated by make and break operations when complex electro-mechanical coupling occurs. The signals can be electrical, acoustic, or optic. Briefly, there are two categories of this approach.

The methods of the first category extract physical characteristics from the acquired data and construct models that map the characteristics to the electrical endurance or contact erosion. Sun *et al.* [5] presented a method that represents the comprehensive evaluation indexes of the electrical life from the probability-weighted means of the closing voltage and the breaking arc energy of every 1000 operations by Spearman ran correlation and the weighted Euclidean distance, and finally mapped the comprehensive evaluation index to the electrical life via unitary regression modeling. Mohammadhosein *et al.* [6] investigated the relationship between eroded mass and characteristics such as the transferred electrical charge, the arc current, the arc energy, and so on for SF_6 circuit breakers and presented an online monitoring scheme. Yoshida *et al.* [7] investigated the influences of multiple characteristics, such as the arc duration time, the bounce time, the arc energy, and the contact resistance, on the electrode mass change of electromagnetic dc contactors, and thus estimated the electrical endurance. However, the methods of this category only work in the circumstance that the entire operation lifetime of an apparatus is monitored, since characteristics reflecting the endurance need to be calculated statistically with a big number of operations or to be compared with the initial status. In other words, they can neither be applied to a used apparatus without historical data nor achieve real-time measurement or prediction.

The methods of the second category do not explore characteristics from the physical point of view, but extract the representative information relating to electrical endurance by machine learning. The methods of this category aim to model the instant working information to reflect the electrical endurance, inspired by the recent studies that applied deep learning to perform fault detection and diagnosis in multiple areas, for example, vehicle gearboxes [8], turbo engines [9], bearings [10], chemical factories [11], manufacturing systems [12], and so on. Wu *et al.* [10] modeled the relationship between the degradation and multisensory signals of bearings by a long short-term memory (LSTM) autoencoder (AE). Chen *et al.* [13] discretized the health status of turbo engines by the enhanced fuzzy c-means algorithm and employed a bi-LSTM network to achieve the remaining useful life (RUL) prediction interval (PI). Li *et al.* [14] proposed an RUL prediction method employing LSTM encoding and a temporal convolutional network (TCN) for turbo engines to improve the prediction precision and efficiency.

Unlike long-term, cyclically operating devices such as engines and bearings, electrical apparatuses perform instantaneous operations. Therefore, the life conditions are reflected by the signals of the transient process, and the machine learning modeling methods are different from those for cyclically

operating devices. Meanwhile, the modeling difficulty is raised due to the complexity of the transient characteristics as well as the constraints of practical purposes. Žarković and Stojković [15] presented an artificial intelligence (AI)-based method for SF_6 circuit breaker health diagnosis. They built three fuzzy expert systems indicating the state of the contacts, the arc extinguishing fluid, and the drive mechanism, and applied k-means clustering followed by artificial neural networks to achieve end-life prediction. The method worked effectively, but acquiring the working information was complicated and costly. Hence, the method can be applied to high-value HV circuit breakers, but is not suitable for frequent switching, low-pricing ac contactors. You *et al.* [16] proposed a fuzzy neural network that utilized acoustic signals caused by mechanical vibrations in make and break operations to predict electrical endurance, based on the phenomenon that contact condition variation changed the acoustic amplitude and spectrum of structural vibrations. This method works efficiently in the laboratory environment, but there are difficulties for industrial applications due to the strong noise interference in the industrial environment.

This article proposes a contacting alloy erosion measurement method that uses the electrical signal waveforms acquired during make and break operations to train a CNNR model that estimates the corresponding contact condition of every single operation. There are two advantages to this method: first, the signals, that is, the voltage and current, are easy to acquire and interference-free; second, the method can be applied online and obtain real-time measurement without historical knowledge.

B. Related Work of the Research Group

The work of this article evolved from the prior research of our research group. The initial study [17] applied conditional density estimation for residual electrical endurance prediction for ac contactors, based on the investigation of the relationship between electrical endurance and breaking arc characteristics. A later paper [18] first proposed to map the break operation signals of ac contactors to several stages of contact mass loss by deep learning classification. Another conference paper [19] preliminarily discussed the feasibility of using CNNR for continuous contact mass loss estimation.

III. CNNR-Based Measurement

A. Overview

In the beginning, the waveforms of the operations are investigated. Fig. 3 shows a make-to-break process. It can be observed that unnormal fluctuations are caused by the instantaneous contacting and multiple bounces at the make operation, as well as the instantaneous separation and the arc discharge at the break operation. Previous studies have proved that the behaviors of the operations change with the variation of the contact condition [20] and also influence the patterns of the signal fluctuations. Hence, we propose an erosion measurement method that achieves the accumulated contact mass loss (ACML) by analyzing the electrical signal fluctuation waveforms acquired during make and break operations. Since most mass loss occurs on the contacting alloy, it is reasonable to represent the erosion condition by the ACML. This approach is a practical solution since acquiring electrical signals is convenient and low cost.

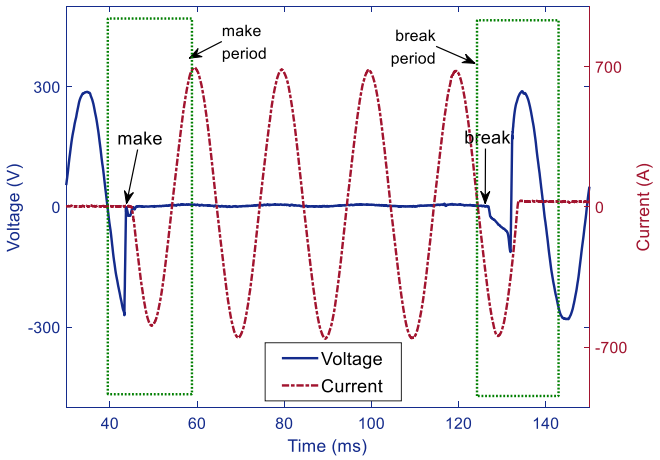

Fig. 3. Fluctuations caused by operations.

Nevertheless, there is a key issue to be resolved: under the ac condition, the initial phase angle of an operation, that is, the contacting moment of two contacts in a make operation or the separation moment of a break operation, remarkably influences the characteristics of the waveforms. It causes difficulties in determining waveform-extracted characteristics that directly reflect the condition of the contacting alloy. To resolve the issue, our research employs deep-learning-based modeling methods that automatically extract representative characteristics from operation waveforms. This article presents an optimized architecture and modeling schemes of the CNNR approach. Comparisons between the architectures and schemes are made to obtain the best measurement result.

B. Data Acquisition and Preprocessing

The research employed the same testing system as what was described in our previous paper [17]. The system runs an ac-4 test by timely performing making and breaking operations on three-phase ac contactors and the amplitude power facture of the current in the main circuit can be configured by setting up the loads. We install a voltage sensor between the input and output terminals of each phase to monitor the voltage variation during operations. Meanwhile, the current of each phase is measured by a current transformer on the wire. All voltage and current waveforms are acquired and recorded by a data acquisition (DAQ).

On the other hand, the contact erosion of a phase is represented by the ACML of the corresponding four contacts (see Fig. 2). The same as described in [18], the corresponding ACML value of an operation is achieved by a piecewise linear interpolation (1) of mass measurements after every 600 operations. M_i represents the corresponding ACML of the ith operation, j is the count of the latest mass-measured operation before the ith operation, and k is the count of the

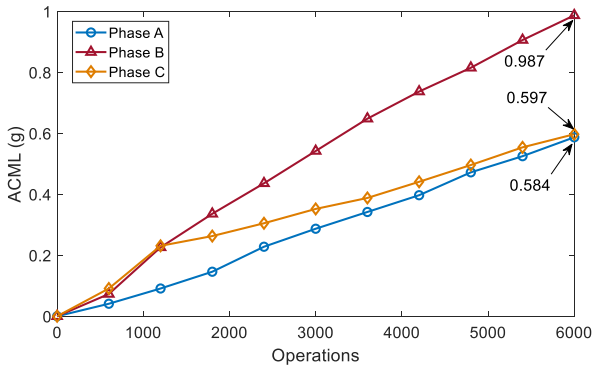

Fig. 4. ACML increment of three-phase contacts.

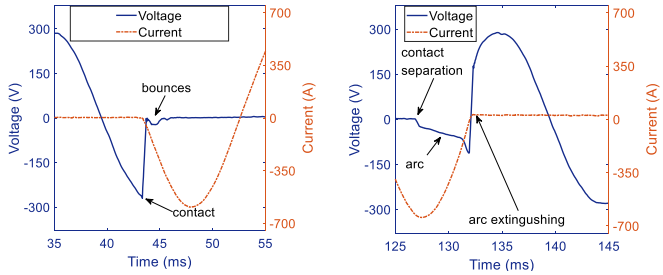

Fig. 5. Waveforms of a make (left) and a break (right).

next mass-measured operation after the ith operation

$$M_i = \frac{(M_k - M_j) \cdot (i - j)}{k - j}. \tag{1}$$

Fig. 4 indicates the ACML increment with the operation count of the three-phase contacts in Fig. 2. It can be observed that the increment of each phase is approximately linear. It also shows an important phenomenon that the erosion rates can be different between phases even in a single ac contactor with uniformly distributed random operation phase angles. The reason is that in break operations, the residual magnet in the driving coil shifts the breaking phase angles close to a certain position and changes the uniformity of breaking phase angles. This phenomenon causes asymmetrical erosion on the contacting alloy of the three phases. The issue was described in our earlier paper [21] in detail. Hence, it also proves that estimating the electrical endurance by counting operations is not reliable and robust.

Hence, the goal of the measurement method is to map the waveform(s) generated by breaking or making an operation to the corresponding ACML which reflects the contact condition. In the previous work, only breaking waveforms are investigated. While in this article, waveforms of making processes are also applied. The length of both breaking and making waveforms is an electricity period. Figs. 3 and 5 show the waveform examples in making and breaking processes.

C. CNNR Architectures

In our previous work, traditional CNNs that performed convolution operations over a space of two or three dimensions were employed, since CNNs were originally designed for applications of images or videos. However, in our research, signals are initially 1-D waveforms. It is considered that reforming 1-D waveforms to a higher dimensionality will ruin the natural features or correlations within or between them. Therefore, a novel architecture that develops our previously proposed network [19] by replacing 2-D convolution layers with 1-D layers is designed. In addition, the novel architecture can also reduce computational complexity. Therefore, 1-D arrays, each of which is an electrical signal waveform, are applied as input features. Comparisons between 1-D and 2-D architectures will be made in Section IV to find out the superior one.

The upper graph of Fig. 6 illustrates the proposed CNN architecture that consists of six 1-D convolution ("Conv") layers, three max-pooling ("Pool") layers, and two full connection ("FC1" and "Output") layers. The height of the convolution and pooling layers indicates the array length of every example, while the width indicates the number of channels in each layer. Since 1-D arrays are used for features, the depth values are all "1." Every convolution layer is followed by a batch normalization layer [22] as well as an activation layer, which are indicated by the deep yellow bands on the right side of each convolution block. The activation function $\xi(x)$ can be rectified linear units (ReLU) [23] or Swish [24]. At the final stage, two full-connection layers ("FC1" and "Output") are designed with a single-value output which is the final measurement result. Unlike classic deep learning networks, the regression network does not employ an activation layer such as SoftMax at the end. The core of the architecture is using three "conv-conv-pool" blocks that increase the network depth and keep the number of feature dimensions. Considering a 1-D convolution layer can be represented as follows:

$$z_u^{(l)} = \sum_{i=-\infty}^{\infty} x_{i+u}^{(l-1)} \cdot k_i^{(l)} \cdot \chi(i) + b^{(l)} \tag{2}$$

where $x^{(l-1)}$ is the initial input of layer l, and $k^{(l)}$ and $b^{(l)}$ are the convolution kernel and the bias, respectively. χ is the valid convolution mask with a mask length n

$$\chi(i) = \begin{cases} 1, & 0 \le i < n \\ 0, & \text{others.} \end{cases} \tag{3}$$

The activation layer is

$$a_u^{(l)} = \xi\left(z_u^{(l)}\right). \tag{4}$$

For two series 1-D convolution layers in a "conv-conv-pool" block, we get

$$z_v^{(l+1)} = \sum_{i=-\infty}^{\infty} a_{i+v}^{(l)} \cdot k_i^{(l+1)} \cdot \chi(i) + b^{(l+1)} \tag{5}$$

$$a_v^{(l+1)} = \xi\left(z_v^{(l+1)}\right). \tag{6}$$

Hence, for the 1-D max-pooling layer with the pooling length r, we get

$$z_u^{(l+2)} = \max\left(a_{ur}^{(l+1)}, a_{ur+1}^{(l+1)}, a_{ur+2}^{(l+1)}, \ldots, a_{ur+r-1}^{(l+1)}\right). \tag{7}$$

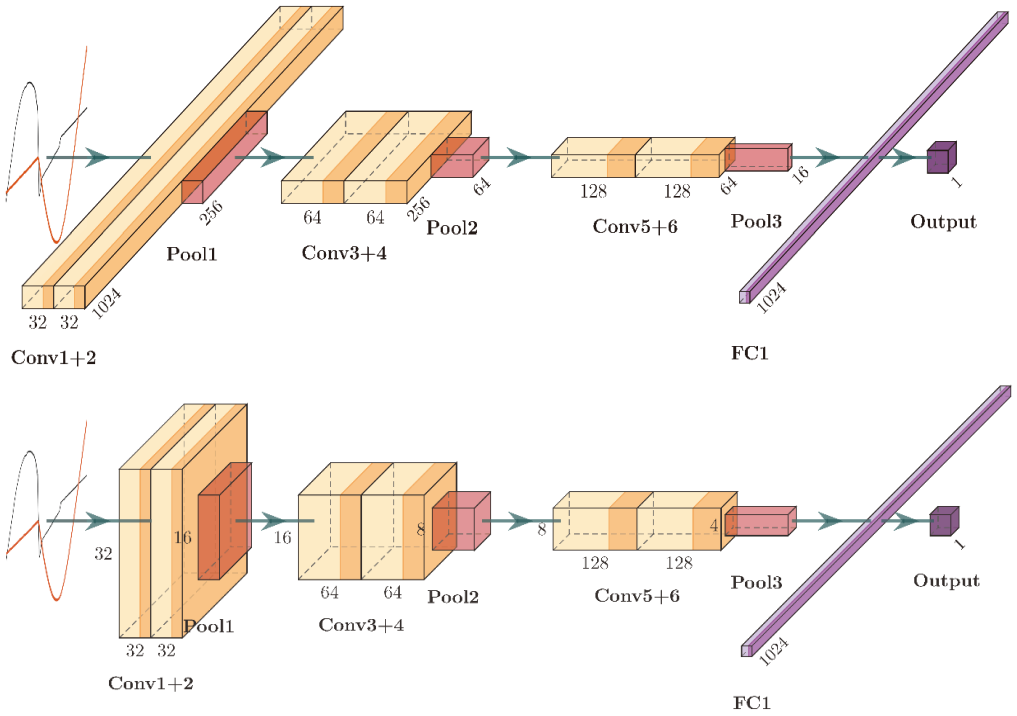

Fig. 6. CNNR architectures: 1-D architecture (upper graph) and 2-D architecture (lower graph).

TABLE I
SPECIFICATIONS OF THE DEVICES USED IN THE EXPERIMENTS (UNIT: g)

Content	Scheme
Operation	Break Make Duo-operation
Channel	Voltage standalone Current standalone Duo-channel
Optimization method	Adam
Batch size	100
Max. epochs	10,000
Dropout rate	0.00 ~ 0.50
Loss function	MSE MAE
Validation Metric	MAE
Data separation	Train : validate : test = 3:1:1
Kernel size	1×25 for 1D 5×5 for 2D

We also investigate a 2-D CNNR architecture that replaces the 1-D convolution layers with 2-D ones (Fig. 6, lower graph), and a reshaping operation that converts 1-D signal arrays to the 2-D matrix is required before inputting data to the network. Meanwhile, 2-D convolution increases the computational complexity and thus reduces the processing efficiency. A comparison between the two architectures is made in Section IV.

D. Training Configuration

We propose several training schemes for comparison (see Table I). We input waveforms occur in make and break operations separately. For each operation, we use voltage standalone, current standalone, as well as duo-channel (i.e., use both voltage and current signals) waveforms. For the duo-channel scheme, voltage and current signals are formed into two channels. We apply both mean squared error (MSE) and mean absolute error (MAE) as the loss function, while only MAE is applied for the validation metric (VM) and test metric (TsM) since it can directly reflect the measurement precision. The training, validation, and test data are randomly divided to 3:1:1 among the entire dataset. Validation is performed with the training running. The model parameters are iteratively replaced according to the progress of the validation precision. The maximum training epoch value is set to 10 000. For the 1-D case, every convolution kernel is a 1×25 1-D array, corresponding to the 5×5 kernel shape commonly used in the 2-D architecture. In our experiments, we have also examined kernels of 1×9, 1×49 for 1-D, as well as 3×3, 7×7 for 2-D, and finally find out that 1×25 and 5×5 are the optimum, respectively. The Adam optimizer is selected since it can adaptively adjust the learning rate and avoid massive parameter tuning. The dropout rate within [0.00, 0.50] in the FC1 layer is discussed, since normally the dropout rate should not be larger than 0.50 in deep learning model training.

The test data is independent of the training and validation data. The final TsM results are obtained by applying the model parameters that achieve the best VM value, that is, the minimum validation MAE, to the test data.

IV. EXPERIMENTS

We train and test CNNR models with TensorFlow 2 [25] and NVIDIA Titan V GPU acceleration cards. The dataset contains 17 925 examples corresponding to the contact mass loss process shown in Fig. 4.

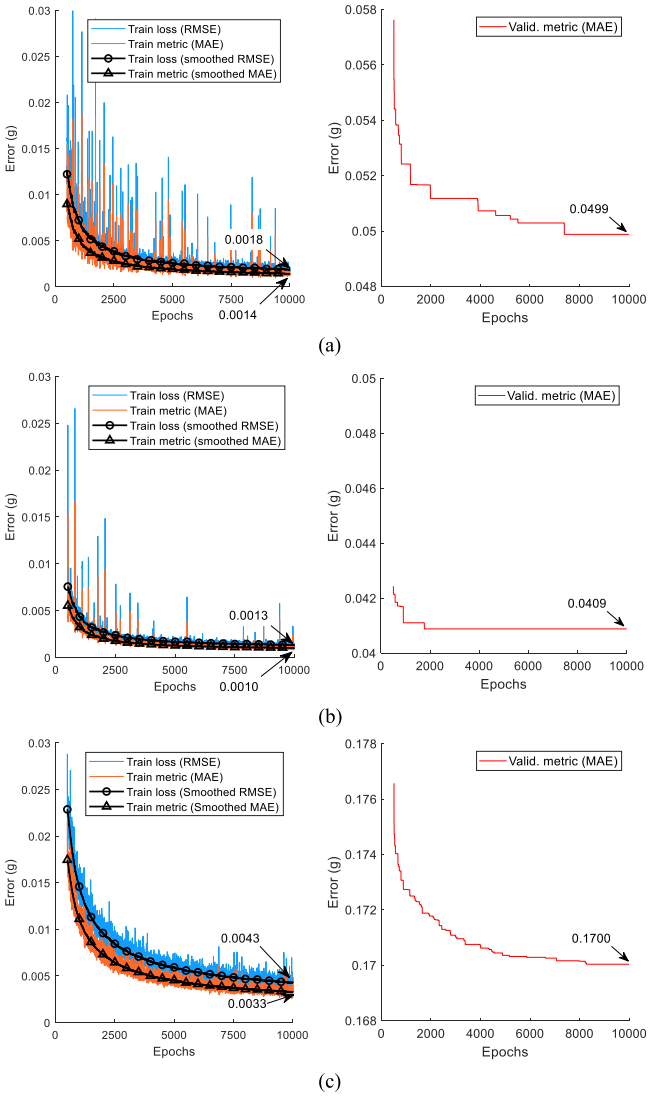

Fig. 7. Training processes of different operations and signals (a) using break-voltage data: left–TrL and TrM; right–VM, (b) using make-voltage data: left–TrL and TrM; right–VM, and (c) using break-current data: left–TrL and TrM; right–VM.

A. Waveforms From Different Operations and Signals

First, we compare the schemes of using waveforms of different operations (make versus break) and different signals (voltage versus current versus duo-channel). In the experiment, the dropout rate equals zero and the loss function is MSE. Only the 1-D CNNR is tested here. Fig. 7 shows three typical examples of the training processes. The loss is converted to the root mean square error (RMSE) from MSE so that the unit of the loss and the metric (MAE) can be unified in the charts.

In Fig. 7, the charts in the left column show the training loss (TrL) and metric values (the first 500 epochs are not presented since the errors at the beginning are very high). Meanwhile, the updating processes of the minimum VM MAEs are displayed in the right column. The "train loss" and "train metric" are smoothed by a sliding mean window of length 100 to form curves (black curves with circular and triangular markers) so that the trend can be easily observed. Obviously, the errors fall quickly and reach the minimum value at the end with the make-voltage waveforms, as indicated in Fig. 7(a). A similar result is obtained by applying the break-voltage data [see Fig. 7(b)]. Nevertheless, Fig. 7(c) shows that the error is significantly high when the break-current data is applied.

TABLE II
TEST RESULTS OF DIFFERENT OPERATIONS AND SIGNALS (UNIT: g)

Opt.	Channel	TrL (RMSE)	TrM (MAE)	VM (MAE)	TsM (MAE)
make	voltage	0.0022	0.0016	0.0409	0.0407
	current	0.0052	0.0043	0.1764	0.1782
	duo	0.0018	0.0015	0.0433	0.0427
break	voltage	0.0030	0.0022	0.0499	0.0494
	current	0.0043	0.0034	0.1700	0.1662
	duo	0.0019	0.0014	0.0541	0.0528
duo	voltage	0.0033	0.0028	0.0504	0.0481
	current	0.0041	0.0035	0.1689	0.1655
	duo	0.0021	0.0015	0.0473	0.0457

TABLE III
TEST RESULTS BY THE 2-D CNNR ARCHITECTURE (UNIT: g)

Opt.	Channel	TrL (RMSE)	TrM (MAE)	VM (MAE)	TsM (MAE)
make	voltage	0.0031	0.0026	0.0618	0.0617
	current	0.0049	0.0040	0.1734	0.1711
	duo	0.0030	0.0024	0.0597	0.0574
break	voltage	0.0020	0.0013	0.0469	0.0489
	current	0.0050	0.0044	0.1513	0.1528
	duo	0.0023	0.0016	0.0476	0.0451
duo	voltage	0.0033	0.0029	0.0507	0.0531
	current	0.0047	0.0043	0.1755	0.1805
	duo	0.0028	0.0020	0.0488	0.0437

Table II shows the TrL, training metric (TrM), VM, and test metric (TsM) at the epoch where the minimum VM MAE is reached. In the column titled "opt." (operation), "duo" means using both make and break waveforms. While in the column titled "channel," "duo" stands for using both voltage and current waveforms. It can be observed that using the make-voltage data reaches the minimum metric errors for both validation and testing. Meanwhile, obviously current waveforms produce negative effects on error reduction, so using duo-channel data increases the computation complexity without precision improvement. Therefore, from the perspective of data, voltage waveforms in make operations can strongly reflect the characteristics of contact mass loss. In other words, from the perspective of electrical and dynamics, the variation of the contact condition changes the bounce behavior in make operations. Table II also shows that the corresponding validation and test MAEs are very close. Considering that the validation and test datasets are randomly splinted in every training process, the generalization of the model can be proven.

B. 1-D and 2-D CNNR Architectures

We also apply the data in the same forms as Section IV-A to our previously proposed 2-D CNNR architecture. The loss and the metric values are shown in Table III. It is indicated that briefly the errors are higher than those achieved by the

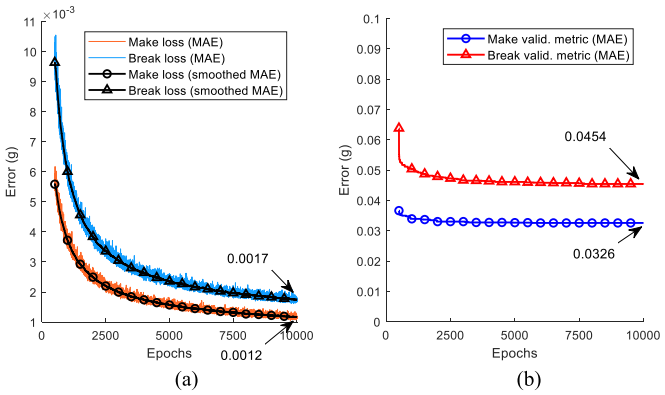

Fig. 8. Training processes using the MAE loss function: (a) TrL. (b) VM.

TABLE IV
TEST RESULTS OF MAE LOSS FUNCTION (UNIT: g)

Configuration			
architecture	1D CNNR		
channel	voltage standalone		
loss function	MAE		
dropout	0.00		
Test results (unit: g)			
operation	TrL=TrM	VM	TsM
make	0.0015	0.0326	0.0324
break	0.0024	0.0454	0.0460

1-D architecture. In addition, the computation cost of the 2-D CNNR is also higher. Therefore, the 1-D CNNR has superiorities in both measurement precision and computational efficiency.

In addition, we also compared the measurement times between the 1-D and 2-D architectures. The average measurement time per example with the 1-D CNNR is 32 μs (including the time of waveform preprocessing), while that with the 2-D CNNR is 37 μs, due to the 2-D computation complexity and the extra process of reshaping feature vectors. In brief, the efficiencies of both 1-D and 2-D architectures can meet the requirement of real-time measurement, owning to the powerful GPU acceleration.

C. Training Parameters

Two important training parameters are investigated: the loss function and the dropout rate. Here, we take the make-voltage data scheme which performs the best in the previous experiments to determine the optimum parameters.

The loss function can also be MAE so that the optimization target of the loss and the metric can be unified. The results are shown in Fig. 8 and Table IV. It is proved that there is a significant improvement in both VM and TsM when employing MAE as the loss function, especially the make-voltage waveforms being used. In addition, the training processes are smoother and more efficient than those applying MSE as the loss function.

TABLE V
TEST RESULTS OF DIFFERENT DROPOUT RATES (UNIT: g)

Configuration			
architecture	1D-CNNR		
channel	voltage standalone		
operation	make		
loss function	MAE		
Test results (unit: g)			
dropout rate	TrL=TrM	VM	TsM
0.00	0.0015	0.0326	0.0324
0.10	0.0111	0.0377	0.0373
0.20	0.0143	0.0333	0.0332
0.30	0.0160	0.0388	0.0377
0.40	0.0167	0.0411	0.0378
0.50	0.0180	0.0443	0.0388

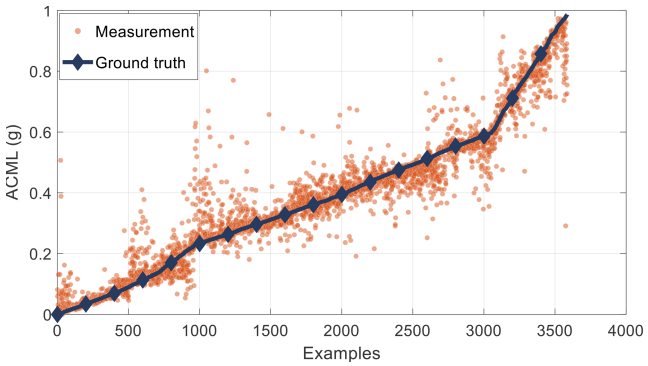

Fig. 9. ACML measurement versus ground truth.

Table V illustrates the results of applying dropout rates from 0.00 to 0.50 in the FC1 layer during training. It is proved that changing the dropout rate cannot improve the precision in our case. Therefore, finally, we achieve an optimum modeling scheme that employs the 1-D CNNR architecture, make-voltage waveforms, the MAE loss function, and the dropout rate of 0.00. The scheme achieves an MAE of 0.0326 g for the VM and an MAE of 0.0324 g for the TsM. In other words, considering the maximum ACML of a phase is 0.984 g, the error rate is 3.29% in absolute values.

In addition, we have also applied Swish as an alternative activation function and achieve the optimal result with TrL = TrM = 0.0087, VM = 0.0448, and TsM = 0.0452. According to the results shown in Table IV, ReLU still reaches a higher precision in our case. The corresponding measurement versus ground-truth chart is displayed in Fig. 9, in which the test examples are sorted by their ground-truth ACML values. It can be observed that most measurement dots tightly surround the ground truth.

D. Comparison Against Other Regression Methods

We implement several existing regression methods for comparative study, including Gaussian process regression (GPR) [26], kernel ridge regression (KRR) [27], support vector regression (SVR) [28], ν-SVR (νSVR) [29], [30],

TABLE VI
COMPARISON WITH OTHER METHODS USING VOLTAGE DATA

Regression method	Test precision (MAE, g)		Avg. time per example (μs)
	make	break	
GPR	0.1083	0.0805	10319
BRR	0.1950	0.1936	13
SGDR	0.1941	0.1925	25
KRR	0.1012	0.0801	246
SVR	0.1017	0.0796	5810
νSVR	0.0974	0.0755	13530
DNNR	0.0812	0.0956	18
LeNet-5 1D-CNNR	0.0714	0.0782	20
LSTM-based	0.0427	0.0566	73
Proposed 1D-CNNR	0.0324	0.0460	32

TABLE VII
RESULTS OF APPLYING THE DATASET OF THE TENGEN TGC1-1811

Regression method	Test precision (MAE, g)		Avg. time per example (μs)
	make	break	
GPR	0.0613	0.0525	8983
BRR	0.1216	0.1133	15
SGDR	0.1229	0.1207	27
KRR	0.1347	0.1209	217
SVR	0.0725	0.0688	21588
νSVR	0.0722	0.0704	11520
DNNR	0.0178	0.0194	19
LeNet-5 1D-CNNR	0.0107	0.0123	20
LSTM-based	0.0405	0.0459	75
Proposed 1D-CNNR	0.0052	0.0092	31

Bayesian ridge regression (BRR) [31], stochastic gradient descent regression (SGDR) [32], deep neural network regression (DNNR) [33], and LeNet-5-like [34] 1-D-CNNR with ReLU activation and max-pooling. The KRR, νSVR, and SVR utilize the radial-basis function (RBF) kernel [26] which obtained higher precision than other kernels, due to the nonlinearity of the regression problem. The DNNR employs a network architecture altered from our 1-D-CNNR architecture by replacing the convolution layers with full connection layers and is trained by the same mechanism as our 1-D-CNNR. GPR, KRR, SVR, νSVR, BRR, and SGDR are implemented with scikit-learn [35], while DNNR and LeNet-5-like 1-D-CNNR are implemented with TensorFlow 2. Also, we implemented a regression model based on LSTM which is commonly applied for RUL prediction. Nevertheless, the supervised LSTM is applied since unsupervised learning employed in [10], [13], and [14] does not work in our experiment. We apply the methods above to both make-voltage and break-voltage data. The parameters of all methods applied are tuned to achieve the best test precision in MAE, and thus the test results are presented in Table VI.

According to Table VI, obviously BRR and SGDR cannot make the precision since they are linear modeling methods. GPR, KRR, νSVR, SVR, DNNR, and LSTM performed better due to their nonlinearity components. Finally, it can be observed that the proposed CNNR reaches significantly higher precision than all other methods. We also present the average measurement time per example in the right column. Since the deep-learning-based methods are accelerated by GPUs, the times are within 100 μs and obviously meet the requirement of contact erosion measurement. Other methods of high computational complexity, such as GPR, SVR, and νSVR, are very time-consuming since GPU acceleration is not supported.

To further verify the proposed method, we also apply another dataset acquired from the ac-4 test of the Tengen TGC1-1811 model whose rated current is 18 A. As the manufacturer requests, the ac-4 current is raised by 20%, that is, 130 A, for a redundancy test. 7200 operations are performed and finally 21 225 valid examples are obtained. At the end of the test, the ACMLs of the three phases are 0.328, 0.298, and 0.184 g, respectively. We apply the same preprocessing, training, and testing schemes to this dataset, and Table VII presents the results of the experiment.

It can be observed that the proposed method also obtains the highest measurement precision (0.0052 g in MSE, i.e., 1.59% of the maximum ACML) as well as a relatively high processing efficiency (31 μs per example).

V. CONCLUSION

This article presents a CNNR method to indirectly measure the contact erosion of ac contactors. Architectural and parametric optimizations are investigated to achieve the best precision. After the comparison with other existing regression methods, it can be concluded that the proposed CNNR method is the optimum approach for the erosion measurement problem at the state of the art.

The work of this article only discusses the measurement with ac-4 test data. However, the load condition also influent the waveforms in operations, and a load of an ac contactor normally varies in practice. It also causes difficulty in applying models to practical applications. The solution is to train a set of models with the data of several typical load conditions, respectively. For a working ac contactor, when an operation is performed under one of the typical load conditions, the corresponding model can be used to achieve the measurement. In further work, data from different load conditions (e.g., ac-3 tests) are expected to be acquired and modeled. Furthermore, multimodel-based contact erosion measurement is to be implemented for varying load conditions. Meanwhile, we will also contribute to developing a device prototype that performs DAQ and contact erosion measurement.

ACKNOWLEDGMENT

The authors would like to thank the People's Government of Yueqing, as well as the Yueqing Municipal Science and Technology Bureau, who support this study by providing research facilities.

REFERENCES

[1] *Low-Voltage Switchgear and Controlgear—Part 4–1: Contactors and Motor-Starters—Electromechanical Contactors and Motor-Starters*, Standard IEC 60947-4-1:2018, International Electrotechnical Commission, Geneva, Switzerland, 2018.

[2] *Low-Voltage Switchgear and Controlgear—Part 2: Circuit-Breakers*, Standard IEC 60947-2:2016, International Electrotechnical Commission, Geneva, Switzerland, 2016.

[3] S. Zheng, F. Niu, K. Li, S. Huang, Z. Liu, and Y. Wu, "Analysis of electrical life distribution characteristics of AC contactor based on performance degradation," *IEEE Trans. Compon., Packag., Manuf. Technol.*, vol. 8, no. 9, pp. 1604–1613, Sep. 2018.

[4] K. Li *et al.*, "Electrical performance degradation model and residual electrical life prediction for AC contactor," *IEEE Trans. Compon., Packag., Manuf. Technol.*, vol. 10, no. 3, pp. 400–417, Mar. 2020.

[5] S. Sun, Q. Wang, T. Du, J. Wang, S. Li, and J. Zong, "Quantitative evaluation of electrical life of AC contactor based on initial characteristic parameters," *IEEE Trans. Instrum. Meas.*, vol. 70, pp. 1–10, 2021.

[6] M. Mohammadhosein, K. Niayesh, A. A. Shayegani-Akmal, and H. Mohsseni, "Online assessment of contact erosion in high voltage gas circuit breakers based on different physical quantities," *IEEE Trans. Power Del.*, vol. 34, no. 2, pp. 580–587, Apr. 2019.

[7] K. Yoshida, S. Shimotsuma, K. Sawa, K. Suzuki, and K. Takaya, "Various characteristics of electromagnetic contactor when arc discharge are generated only make arc," in *Proc. IEEE 62nd Holm Conf. Electr. Contacts (Holm)*, Oct. 2016, pp. 215–221.

[8] Y. Liao, X. Zeng, and W. Li, "Wavelet transform based convolutional neural network for gearbox fault classification," in *Proc. Prognostics Syst. Health Manage. Conf. (PHM-Harbin)*, Jul. 2017, pp. 1–6.

[9] M. Hou, D. Pi, and B. Li, "Similarity-based deep learning approach for remaining useful life prediction," *Measurement*, vol. 159, Jul. 2020, Art. no. 107788.

[10] J.-Y. Wu, M. Wu, Z. Chen, X.-L. Li, and R. Yan, "Degradation-aware remaining useful life prediction with LSTM autoencoder," *IEEE Trans. Instrum. Meas.*, vol. 70, pp. 1–10, 2021.

[11] W. Yu and C. Zhao, "Broad convolutional neural network based industrial process fault diagnosis with incremental learning capability," *IEEE Trans. Ind. Electron.*, vol. 67, no. 6, pp. 5081–5091, Jun. 2020.

[12] L. Wen, X. Li, L. Gao, and Y. Zhang, "A new convolutional neural network-based data-driven fault diagnosis method," *IEEE Trans. Ind. Electron.*, vol. 65, no. 7, pp. 5990–5998, Jul. 2018.

[13] C. Chen, N. Lu, B. Jiang, Y. Xing, and Z. H. Zhu, "Prediction interval estimation of aeroengine remaining useful life based on bidirectional long short-term memory network," *IEEE Trans. Instrum. Meas.*, vol. 70, pp. 1–13, 2021.

[14] J. Li, R. Chen, and X. Huang, "A sequence-to-sequence remaining useful life prediction method combining unsupervised LSTM encoding-decoding and temporal convolutional network," *Meas. Sci. Technol.*, vol. 33, no. 8, Aug. 2022, Art. no. 085013.

[15] M. Žarković and Z. Stojković, "Artificial intelligence SF_6 circuit breaker health assessment," *Electr. Power Syst. Res.*, vol. 175, Oct. 2019, Art. no. 105912.

[16] Y. You, J. Wang, G. Wu, L. Shu, and Z. Wu, "Electrical life prediction method of AC contactor based on fuzzy neural network," in *Proc. 7th Int. Conf. Rel. Electr. Products Electr. Contacts (ICREPEC)*, Suzhou, China, Nov. 2019, pp. 265–269.

[17] Z. Wu, G. Wu, H. Huang, and Y. You, "A novel residual electrical endurance prediction method for low-voltage electromagnetic alternating current contactors," *IEEE Trans. Compon., Packag., Manuf. Technol.*, vol. 5, no. 4, pp. 465–473, Apr. 2015.

[18] H. Cui, Z. Wu, G. Wu, X. Xu, Y. You, and Y. Fang, "Convolutional neural networks for electrical endurance prediction of alternating current contactors," *IEEE Trans. Compon., Packag., Manuf. Technol.*, vol. 9, no. 9, pp. 1785–1793, Sep. 2019.

[19] Z. Wu, H. Cui, G. Wu, and Y. You, "Research on convolutional neural network regression for contact erosion estimation of AC contactors," in *Proc. 7th Int. Conf. Rel. Electr. Products Electr. Contacts (ICREPEC)*, Suzhou, China, Nov. 2019, pp. 45–50.

[20] F. Pons, "Electrical contact material arc erosion: Experiments and modeling towards the design of an AgCdO substitute," Ph.D. thesis, School Mech. Eng., Georgia Inst. Technol., 2010.

[21] Z. Wu, G. Wu, C. Chen, Y. Fang, L. Pan, and H. Huang, "A novel breaking strategy for electrical endurance extension of electromagnetic alternating current contactors," *IEEE Trans. Compon., Packag., Manuf. Technol.*, vol. 6, no. 5, pp. 749–756, May 2016.

[22] V. Thakkar, S. Tewary, and C. Chakraborty, "Batch normalization in convolutional neural networks—A comparative study with CIFAR-10 data," in *Proc. 5th Int. Conf. Emerg. Appl. Inf. Technol. (EAIT)*, Jan. 2018, pp. 1–5.

[23] X. Glorot, A. Bordes, and Y. Bengio, "Deep sparse rectifier neural networks," *J. Mach. Learn. Res.*, vol. 15, pp. 315–323, Jun. 2011.

[24] P. Ramachandran, B. Zoph, and Q. V. Le, "Searching for activation functions," CoRR, *vol. abs/1710.05941*, pp. 1–13, Oct. 2017.

[25] M. Abadi, P. Barham, J. Chen, Z. Chen, and X. Zhang, "TensorFlow: A system for large-scale machine learning," in *Proc. 12th USENIX Symp. Operating Syst. Des. Implement. (OSDI)*, 2016, pp. 265–283.

[26] C. E. Rasmussen and C. K. I. Williams, *Gaussian Processes for Machine Learning*. Cambridge, MA, USA: MIT Press, 2006.

[27] S. Theodoridis, "Learning in reproducing kernel Hilbert spaces," in *Machine Learning*, 2nd ed. New York, NY, USA: Academic, 2020, ch. 11, pp. 531–594.

[28] M. Awad and R. Khanna, "Support vector regression," in *Efficient Learning Machines*. Berkeley, CA, USA: Apress, 2015, pp. 67–80.

[29] J. C. Platt, "Probabilistic outputs for support vector machines and comparisons to regularized likelihood methods," in *Advances in Large-Margin Classifiers*. Cambridge, MA, USA: MIT Press, 2000, pp. 61–74.

[30] C. Chang and C. J. Lin, "Training υ-support vector regression," *Neural Comput.*, vol. 14, no. 8, pp. 1959–1977, Aug. 2002.

[31] Y. Liu, L. Sun, C. Du, and X. Wang, "Near-infrared prediction of edible oil frying times based on Bayesian ridge regression," *Optik*, vol. 218, Sep. 2020, Art. no. 164950.

[32] W. Xu, "Towards optimal one pass large scale learning with averaged stochastic gradient descent," 2011, *arXiv:1107.2490*.

[33] L. Ngo, J. Cha, and J.-H. Han, "Deep neural network regression for automated retinal layer segmentation in optical coherence tomography images," *IEEE Trans. Image Process.*, vol. 29, pp. 303–312, 2020.

[34] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proc. IEEE*, vol. 86, no. 11, pp. 2278–2324, Nov. 1998.

[35] F. Pedregosa *et al.*, "Scikit-learn: Machine learning in Python," *J. Mach. Learn. Res.*, vol. 12, no. 10, pp. 2825–2830, 2011.

Ziran Wu was born in Wenzhou, China, in 1984. He received the Ph.D. degree in electronic engineering from the University of York, York, U.K., in 2013.

Since 2014, he has been with Wenzhou University, Wenzhou, where he is currently an Associate Professor with the Electrical and Electronic Engineering College. He is the Deputy Director of the Technology Institute, Wenzhou University, Yueqing, China. His current research interests include electrical apparatuses, artificial intelligence (AI), and smart manufacturing.

Dr. Wu is also a Committee Member of the International Conference on Reliability of Electrical Products and Electrical Contacts.

Chufu Fang was born in Jiangxi, China, in 1975. He received the B.B.A. degree in business administration from Zhejiang University, Hangzhou, China, in 2010.

Since 1998, he has been with Zhengjiang Tengen Electric Company, Ltd., Wenzhou, China, and he is currently the Vice President in charge of the manufacturing and informalization management of the corporation. He focuses on the planning and construction of smart factories, information systems, and lean manufacturing systems, as well as the optimization of the supply chain.

Guichu Wu was born in Wenzhou, China, in 1957. He received the B.S. degree in physics from the Normal College of Wenzhou, Wenzhou, Zhejiang, China, in 1980.

In 1980, he started work at Wenzhou University, Wenzhou, and retired in 2020. Since 2020, he has been the Chairperson of the board of Zhejiang Juchuang Smartech Company Ltd., Wenzhou. His current research interests include smart electrical apparatuses, automatic and digital manufacturing, embedded system applications, and new energy resources.

Prof. Wu is also the Honorary Director of the Technology Institute, Wenzhou University, Yueqing, China, and a committee member of the International Conference on Reliability of Electrical Products and Electrical Contacts.

Zhenquan Lin was born in Wenzhou, China, in 1963. He received the Ph.D. degree in physics from Zhejiang University, Hangzhou, China, in March 2014.

Since 1988, he has been with Wenzhou University, Wenzhou, where he is currently a Professor with the Mathematics and Physics College. He is the Director of the Technology Institute, Wenzhou University, Yueqing, China. His current research interests include theoretical condensed matter physics and statistical analysis on spin materials.

Prof. Lin is also a Standing Director of the Zhejiang Physical Society.

Wei Chen was born in Wenzhou, China, in 1989. He received the B.S. and Ph.D. degrees in electrical engineering from Zhejiang University, Hangzhou, China, in 2011 and 2016, respectively.

He is currently an Associate Professor with the Technology Institute, Wenzhou University, Yueqing, China, and the College of Electrical and Electronic Engineering, Wenzhou University, Wenzhou, Zhejiang, China. His current research interests include design and analysis of high-speed permanent magnet synchronous machines, high-precision current sensors, and electric vehicles' drive machines.