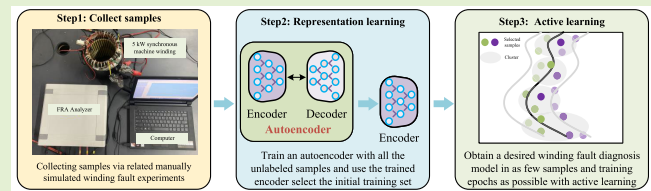


Data-Efficient Synchronous Machine Winding Short-Circuit Faults Diagnosis Based on Frequency Response Analysis and Active Learning

Yu Chen^{ID}, Zhongyong Zhao^{ID}, *Member, IEEE*, Jiangnan Liu^{ID}, Wei Wang^{ID},
and Chenguo Yao^{ID}, *Member, IEEE*

Abstract—Winding short-circuit (SC) faults are a prevalent issue in synchronous machines, and the accurate and timely identification of these faults is critical for maintaining power system stability. Current methods for inspecting synchronous machine windings often rely on periodic inspections based on human expertise. Therefore, numerous studies have explored the application of deep learning (DL) models for detecting synchronous machine winding SC faults. However, these models often exhibit excessive complexity and overlook physical overhead, leading to inefficient utilization of computational power and data resources. To address these limitations, this study proposes a dual-channel DL model integrated with the active learning (AL) query strategy. In this study, winding SC faults are manually simulated on a 5-kVA synchronous machine, and corresponding frequency response analysis (FRA) data are recorded. Subsequently, the proposed method is validated on the test set and benchmarked against previous studies. Experimental results demonstrate that the proposed method significantly reduces the data annotation effort and accelerates model training to the order of seconds (under 5 s) while maintaining satisfactory accuracy ($\geq 95\%$). Comparative experimental results further indicate that the proposed model, requiring only 1/20th of the labeled training samples used in previous studies, achieves a substantial reduction of approximately 99.8% in both model parameters and training time.

Index Terms—Active learning (AL), fault diagnosis, frequency response analysis (FRA), synchronous machine, winding fault.



I. INTRODUCTION

SYNCHRONOUS machines, the core components of hydroelectric and thermal power plants, are primarily employed to convert mechanical energy into electrical energy. A related survey [1] indicates that winding faults constitute approximately two-thirds of all failures in large synchronous machines, with winding short-circuit (SC) faults being the most prevalent type. Winding SC faults are primarily attributed to the degradation of insulation material, a cumulative process that ultimately leads to insulation breakdown. This breakdown can result in various types of winding SC faults, including inter-turn SC (ITSC) and ground SC (GSC) faults. These SC faults not only damage the synchronous machine itself but also significantly impact the entire power system, potentially leading to large-scale power outages. Consequently, the rapid and accurate diagnosis of winding SC faults in synchronous machines is of paramount importance for maintaining the stable operation of power systems and safeguarding the economic interests of associated industries. 9

Received 18 March 2025; accepted 4 April 2025. Date of publication 15 April 2025; date of current version 2 June 2025. This work was supported in part by Sichuan Science and Technology Program under Grant 2023NSFSC0829, in part by the Fundamental Research Funds for the Central Universities under Grant SWU-KT22027, and in part by the National Natural Science Foundation of China under Grant 51807166. The associate editor coordinating the review of this article and approving it for publication was Prof. Qing Ni. (*Corresponding author: Zhongyong Zhao.*)

Yu Chen is with the College of Engineering and Technology, Southwest University, Chongqing 400716, China, and also the Department of Building Environment and Energy Engineering, The Hong Kong Polytechnic University, Hong Kong (e-mail: yu_chen2000@hust.edu.cn).

Zhongyong Zhao is with the College of Engineering and Technology, Southwest University, Chongqing 400716, China, and also with Yibin Academy of Southwest University, Yibin, Sichuan 644000, China (e-mail: zhaozy1988@swu.edu.cn).

Jiangnan Liu is with Guangzhou Power Supply Bureau, Guangdong Power Grid Company Ltd., Guangzhou 510620, China (e-mail: xjy7641@sina.com).

Wei Wang is with the State Key Laboratory of Advanced Electromagnetic Technology, School of Electrical and Electronic Engineering, Huazhong University of Science and Technology, Wuhan 430074, China (e-mail: husterww@hust.edu.cn).

Chenguo Yao is with the State Key Laboratory of Power Transmission Equipment and System Security and New Technology, Chongqing University, Chongqing 400044, China (e-mail: yaochenguo@cqu.edu.cn).

Digital Object Identifier 10.1109/JSEN.2025.3559011

Existing noninvasive condition monitoring techniques for synchronous machines, such as those employing thermal analysis [2] or vibration signal analysis [3], offer real-time operational insights [4]. However, retrofitting existing machines with the necessary sensors is challenging and costly due to their fixed internal structures. Consequently, most fault diagnosis methods necessitate offline testing, requiring machine shutdown. Currently, there are various winding fault diagnostic approaches, including motor current signature analysis (MCSA) [5], recurrent surge oscilloscope (RSO) [6], traveling wave methods [7], and so on.

Frequency response analysis (FRA), a well-established offline diagnostic technique for transformer windings, has been increasingly applied to synchronous machines due to structural similarities [8], [9], [10]. Its rapid and nondestructive nature, combined with high stability and sensitivity, renders it suitable for assessing winding conditions during routine maintenance. FRA-based fault diagnosis commonly utilizes one of three primary approaches: 1) broadband equivalent circuit modeling predicated on winding structure; 2) indicator-based fault detection relying on deviations from baseline FRA signatures; and 3) data-driven fault diagnosis methodologies.

Equivalent circuit models have been widely adopted in several studies for interpreting FRA data. These include a black-box model designed to assess FRA applicability [11], a gray-box model developed for detecting ITSCs, and a high-frequency lumped parameter model based on broadband characteristics [12]. While such modeling methods provide valuable physical insights for expert diagnosis, their inherent complexity and time-intensive development limit rapid and general fault detection, primarily because the diverse designs of synchronous machines necessitate distinct equivalent models.

Numerous studies have proposed indicator-based methods for detecting winding faults using FRA data. These methods typically utilize mathematical–statistical indicators, such as minimum–maximum ratios and normalized correlation coefficients, which exhibit a linear correlation with the severity of the fault [13], [14], [15]. Correlation coefficients have also been employed to detect GSCs and ITSCs [10]. While some researchers have combined these indicators with classifiers to enhance the accuracy of fault detection, the effective application of these indicators often necessitates expert interpretation and may require adjustments contingent upon the specific machine type.

The lack of standardized FRA interpretation codes for synchronous machines necessitates a reliance on expert subjective judgment in fault diagnosis. Data-driven approaches present a promising avenue for facilitating objective assessments [18]. Machine learning (ML) techniques that have been applied to address this challenge include isolation forest [16] and support vector machines (SVMs) for fault detection [17], as well as an explainable artificial intelligence (AI) technology designed to improve model interpretability [18]. However, despite the acknowledged potential of advanced ML and deep learning (DL) models, significant technical challenges persist.

- 1) A common approach involves the direct application of classical DL backbone models for the detection of winding faults. However, these models are often characterized by a substantial number of parameters, necessitating significant computational resources for both training and subsequent deployment.
- 2) Although manual experiments conducted in a laboratory setting provide convenient access to a variety of labeled data, attaining comparable levels of diagnostic accuracy with operational synchronous machines requires a significant amount of workforce and time for data annotation. Furthermore, maintenance personnel can quickly identify the occurrence of winding faults through external measurements, but disassembling faulty machines to ascertain fault types is time-consuming.
- 3) Power plants and equipment manufacturers possess numerous unlabeled data rather than labeled data. However, existing DL-based methods for the winding fault diagnosis currently do not utilize these resources.

Active learning (AL), which strategically selects unlabeled samples for labeling to maximize model performance with the fewest labeled samples [19], [20], [21], [22], offers a potential solution. AL has been applied in power system applications, such as stability assessment [23] and identification of instability modes [20], both of which also have the same challenges. Inspired by the mentioned works [20], [23], this study proposes a lightweight DL model incorporating AL for winding SC fault diagnosis. The main contributions are as follows.

- 1) Unlike the direct application of classical DL models, we propose a dual-channel DL model that significantly reduces parameters while maintaining precision. Furthermore, in contrast to conventional single-channel FRA models, the proposed model is dual-channel, considering both FRA gain and phase data.
- 2) Diverging from previous works that are solely focused on improving model performance, we develop a two-stage batch-mode AL query strategy designed to minimize the data annotation burden by selecting the most representative samples from the unlabeled pool for labeling.
- 3) In contrast to prior methods that do not leverage unlabeled data, we propose a general framework for integrating classification models with AL, which effectively utilizes all available unlabeled data.

The remainder of this study is organized as follows. Section II introduces the fundamentals of FRA, AL, and the proposed AL framework. Section III details the FRA dataset obtained from manually simulated winding SC experiments. Section IV presents the results and compares the proposed method with existing approaches. Section V discusses limitations, and conclusions are provided in Section VI.

II. METHODOLOGY

DL models offer an effective tool for detecting winding SC faults [18], [26]. Fig. 1 shows the overall framework of the proposed dual-channel DL model combined with AL query strategies, as well as its detailed architecture. In Fig. 1, the

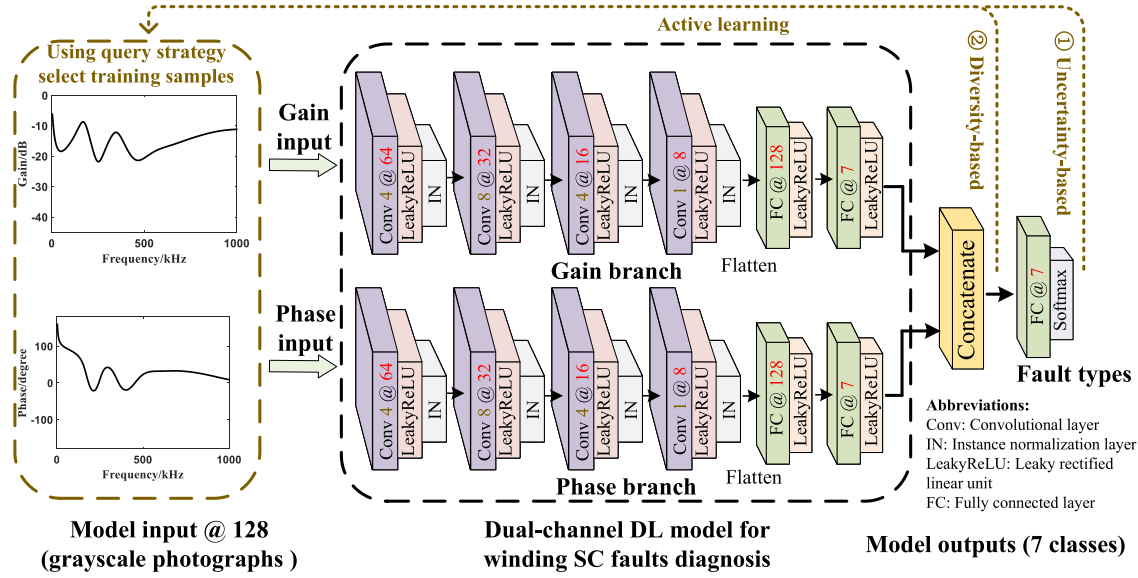


Fig. 1. Overall framework of the dual-channel DL model with the AL query strategy for winding SC fault diagnosis. The “Conv” and “@” are followed by the output channel and size (two dimensions in the same size). The dual-channel DL model and the linear layer output serve as the feature input for diversity- and uncertainty-based query strategies.

model’s inputs consist of two 128×128 grayscale images representing the FRA gain and phase, and the model’s output serves as the feature input for AL query strategies. To provide classification outputs for winding SC fault diagnosis, a linear layer is integrated after the dual-channel DL (backbone) model. In addition, Chen et al. [18] explained the reasons and advantages of selecting an image-classification model.

A. Basic Principle of FRA

Under sweep signal excitation, the synchronous machine winding can be represented by an equivalent circuit model comprising resistance, inductance, and capacitance [24], as shown in Fig. 2(a). The sweep signal is applied as the excitation signal $v_{in}(\omega)$ at one terminal of the winding, while the corresponding response signal $v_{out}(\omega)$ is recorded at the other. Subsequently, the excitation and response signals are used to derive the FRA gain $H(\omega)$ and phase $\Phi(\omega)$, as presented in (1) and (2). These FRA data characterize the condition of the winding, distinguishing between normal and faulty states. The distributed parameters inherently reflect the geometric dimensions and conditions of the winding; consequently, any winding fault alters these parameters, resulting in variations in $H(\omega)$ and $\Phi(\omega)$, as shown in Fig. 2(b). By comparing the measured frequency response curve with a baseline (normal) curve, FRA enables the assessment of the winding’s condition, including the detection of the occurrence, type, severity, and location of winding faults

$$H(\omega) = 20 \log_{10} \left| \frac{v_{out}(\omega)}{v_{in}(\omega)} \right| \text{ dB} \quad (1)$$

$$\phi(\omega) = \frac{180}{\pi} \arg \left| \frac{v_{out}(\omega)}{v_{in}(\omega)} \right| \quad (2)$$

where $v_{out}(\omega)$ and $v_{in}(\omega)$ are the excitation and response signals, respectively, and $H(\omega)$ and $\Phi(\omega)$ are the gain and

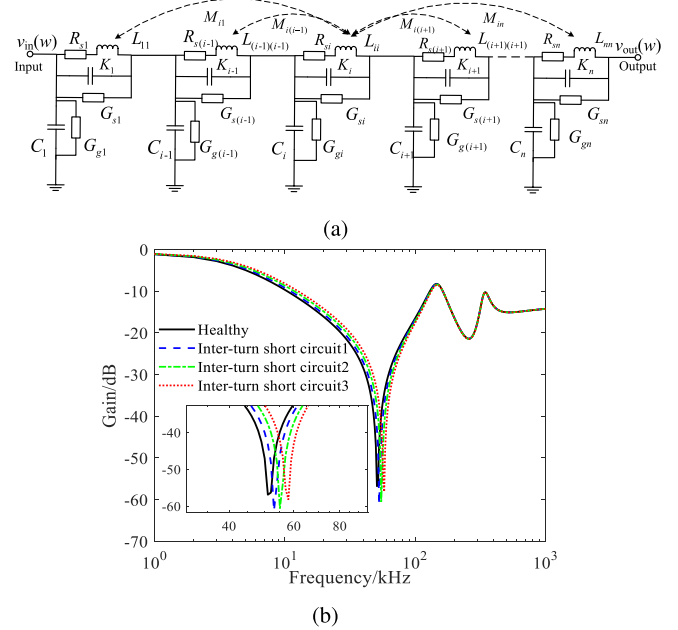


Fig. 2. Synchronous machine winding's equivalent circuit model structure and simulated winding fault. (a) Equivalent circuit model. (b) ITSCs simulated by the built equivalent circuit [24].

phase of the transfer function, respectively. This study uses a frequency range from 1 Hz to 1000 kHz.

B. Basic Principle of AL and Designed Query Strategy

Data-driven models have demonstrated remarkable performance in various tasks [5], [18], [25], [26], whose success heavily relies on high-quality labeled datasets. In the specific context of synchronous machine winding fault diagnosis, equipment manufacturers and power plants often possess a large number of unlabeled samples. The annotation of

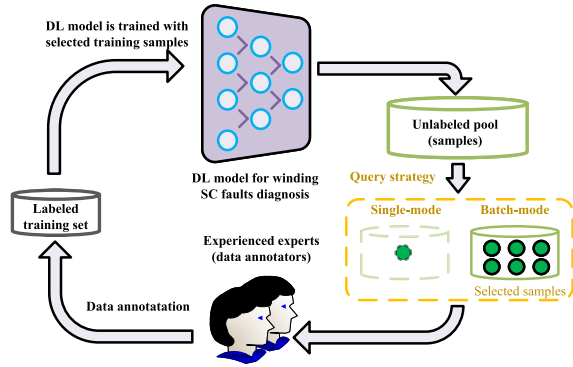


Fig. 3. Pool-based AL cycle with the classification task and comparison between the single-mode and batch-mode AL cycles.

these samples is expensive and time-consuming, making the complete labeling of such datasets impractical. Consequently, the effective utilization of these unlabeled data, through the development of methods that selectively label the most informative samples, is crucial. This capability represents a prerequisite for the practical application of data-driven models to fault diagnosis in this domain.

AL aims to optimize model performance with minimal labeled data, thereby reducing annotation costs. Fig. 3 illustrates the common pool-based AL cycle. This involves an unlabeled dataset $U_n^t = \mathcal{X}, \mathcal{Y}$ (n samples) and a labeled training set $L_m^t = X, Y$ (m samples), where $\mathcal{X}$ and X represent sample spaces and $\mathcal{Y}$ and Y represent label spaces, respectively. The potential distribution is denoted by $\mathbb{P}_{X,Y}$. The variable t represents the number of iterations, and each iteration involves labeling b samples. The goal of this study is to design an effective AL query strategy Q and utilize the learning algorithm $\mathcal{T}$ to train a multiclass classification model $\mathcal{M}_{\theta_t} : X \rightarrow Y$ with as few samples or t as possible. The optimization problem of AL can be defined as follows:

$$\arg \min_{t, (x,y) \in L_m^t, (x,y) \in U_n^t} \mathbb{E}_{x,y \sim \mathbb{P}_{X,Y}} [\ell(\mathcal{M}_{\theta_t}(x), y; \mathcal{T}; Q)] \quad (3)$$

where $\ell(\cdot)$ is the given loss function. The specific procedure is shown in Algorithm 1.

The core of AL lies in the query strategy Q , which selects the most informative samples for training, thereby accelerating model performance improvement. Common strategies can be categorized into uncertainty- and diversity-based approaches [19], [21].

- 1) Uncertainty-based query strategies, a prevalent and straightforward approach, prioritize the selection of samples that exhibit high uncertainty for labeling. This selection process concentrates on the most informative samples, which are inherently characterized by high uncertainty, leading to potential improvements in model performance. Common metrics for quantifying uncertainty include least confidence, margin sampling, and entropy [19]. Specifically, high uncertainty typically manifests as low confidence scores, small margin probabilities, or high entropy values. Least confidence is defined as follows:

$$P_{\text{least}} = \mathbb{P}_{\mathcal{M}_{\theta}}(\hat{y}_1 | x), \quad x \in U \quad (4)$$

Algorithm 1 General Pool-Based AL Framework Combined With Winding SC Faults Diagnosis

Require: Unlabeled dataset U , current training set L , test set D^T , batch size b , maximum iterations T , test set accuracy Acc_{test} , desired accuracy target Acc_{tar} , data-driven model $\mathcal{M}_{\theta}$, learning algorithm $\mathcal{T}$, and AL query strategy Q .

- 1: $t = 1$.
- 2: Randomly select b unlabeled samples and label them to construct U_{n-b}^0 and L_b^0 . // Construct the initial training set. If there are existing labeled samples, they can be used as the initial training set.
- 3: $\mathcal{M}_{\theta_0} \xleftarrow{\mathcal{T}} L_b^0$ // Use initial labeled training set to train the DL model $\mathcal{M}_{\theta_0}$.
- 4: **while** $t \leq T$ and $Acc_{test} < Acc_{tar}$ **do**
- 5: $U_{n-b \times (t+1)}^t, L_{b \times (t+1)}^t \leftarrow Q(U_{n-b \times t}^{t-1}, L_{b \times t}^{t-1}, \mathcal{M}_{\theta_{t-1}})$ // Use query strategy to select the top b unlabeled samples and hand them to experienced experts to label them. The current training set $L_{b \times (t+1)}^t$ consists of the old training set $L_{b \times t}^{t-1}$ and b new labeled samples. Single-mode or batch-mode refers to selecting one or multiple unlabeled samples.
- 6: $\mathcal{M}_{\theta_t} \xleftarrow{\mathcal{T}} L_{b \times (t+1)}^t$ // Use current labeled training set with the learning algorithm to train the data-driven model.
- 7: $Acc_{test} \xleftarrow{D^T} \mathcal{M}_{\theta_t}$ // Validating the trained data-driven model on the test set.
- 8: $t \leftarrow t + 1$
- 9: **end while**
- 10: **return** $\mathcal{M}_{\theta_t}$

where x is the input data from the unlabeled pool, while $\hat{y}_1$ is the most probable class and $\mathbb{P}_{\mathcal{M}_{\theta}}(\hat{y}_1 | x)$ is the posterior probability prediction of $\hat{y}_1$. Margin probability is defined as follows:

$$P_{\text{margin}} = \mathbb{P}_{\mathcal{M}_{\theta}}(\hat{y}_1 | x) - \mathbb{P}_{\mathcal{M}_{\theta}}(\hat{y}_2 | x), \quad x \in U \quad (5)$$

where $\hat{y}_2$ is the second probable class and $\mathbb{P}_{\mathcal{M}_{\theta}}(\hat{y}_2 | x)$ is the posterior probability prediction of $\hat{y}_2$. Entropy is defined as follows:

$$P_{\text{entropy}} = - \sum_i \mathbb{P}_{\mathcal{M}_{\theta}}(\hat{y}_i | x) \log \mathbb{P}_{\mathcal{M}_{\theta}}(\hat{y}_i | x), \quad x \in U \quad (6)$$

where $\hat{y}_i$ ranges over all the probable classes. Although uncertainty-based query strategies are effective in identifying individually informative samples, they frequently neglect to consider the potential joint informativeness among samples, failing to adequately account for potential information redundancy among the selected instances, particularly when those samples exhibit mutual similarity. Consequently, consistent selection of similar samples introduces the risk of overfitting and can increase the difficulty of training. Furthermore, the mentioned formulas (4)–(6) are designed for the conventional single-model AL. In batch-mode AL, the

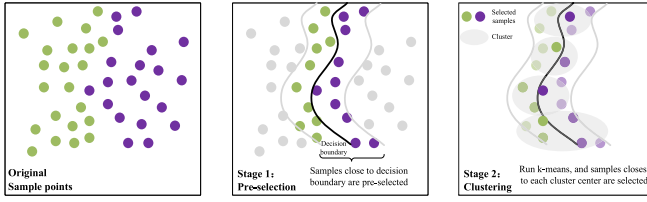


Fig. 4. Flowchart of two-stage batch-mode query strategy.

uncertainty scores of unlabeled samples are computed and ranked, subsequently followed by the selection of the top b samples.

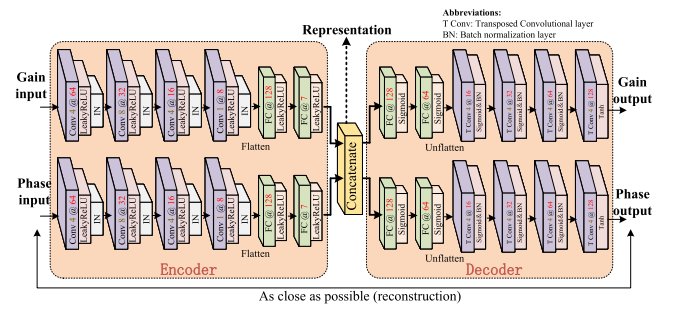
- 2) Diversity-based query strategies effectively capture the distribution of the unlabeled data pool by selecting representative samples. A common approach is cluster-based sampling, which divides the unlabeled data into different groups and selects a representative sample from each group to ensure diversity. K -means is a common clustering algorithm. Given unlabeled samples $\{x_1, x_2, \dots, x_n\} \in U$, the objective of K -means can be defined as followed:

$$\arg \min_S \sum_{i=1}^k \sum_{\mathbf{x} \in S_i} \|\mathbf{x} - \mu_i\|^2 \quad (7)$$

where $S = \{S_1, S_2, \dots, S_k\}$ is the set of clusters, and μ_i is the mean of points in S_k (cluster centers). After determining cluster centers, samples closest to each cluster center are selected and labeled. The diversity-based query strategy aims to encompass the distribution of the unlabeled pool but often prioritizes diversity over proximity to the decision boundary. Therefore, this strategy can lead to the selection of samples far from the decision boundary, which are unlikely to improve model performance due to their low misclassification risk.

- 3) As previously discussed, single-stage query strategies exhibit inherent limitations. Therefore, this study proposes a two-stage batch-mode query strategy that integrates uncertainty and diversity, whose main concept is measuring uncertainty through preselection and then choosing diverse samples through clustering. Specifically, this study introduces two integrated query strategies:

- a) *Integrated Query Strategy 1*: Preselect $r \times b$ samples using margin probability (5), where r ($r \in \mathbb{N}^+, r > 1$) is a hyperparameter to control the preselected sample size. Among the preselected $r \times b$ samples, use K -means (set $k = b$) to select b samples closest to cluster centers. The flowchart is shown in Fig. 4.
- b) *Integrated Query Strategy 2*: The entropy (6) of the unlabeled samples is calculated, which serves as sample weights during fitting clusters of K -means (set $k = b$). Then, select b samples closest to the cluster centers.



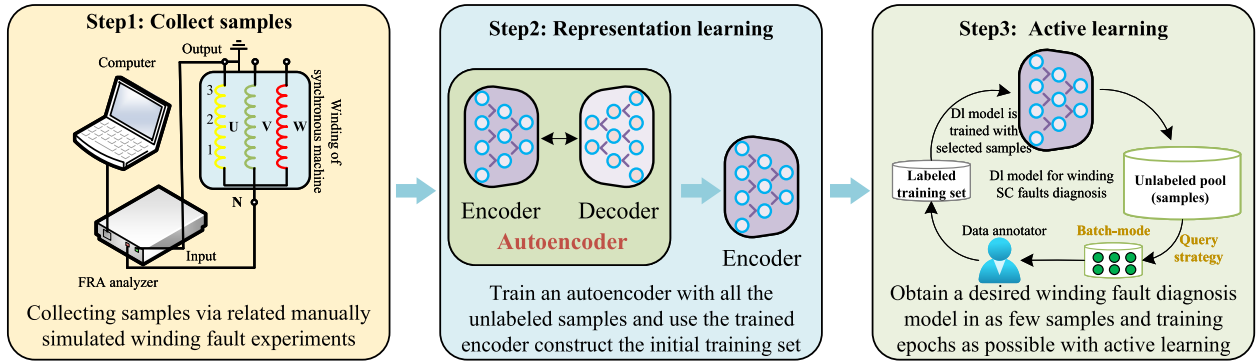


Fig. 6. Flowchart of the proposed method.

Algorithm 2 General AL Framework Proposed for Winding SC Faults Diagnosis

Require: Unlabeled dataset U , current training set L , test set D^T , batch size b , maximum iterations T , test set accuracy Acc_{test} , desired accuracy target Acc_{tar} , encoder ϕ , decoder ψ , the proposed dual-channel DL model $\mathcal{M}_\theta$, and learning algorithm $\mathcal{T}$.

- 1: $t = 1$
- 2: $\phi, \psi \leftarrow \arg \min_{\phi, \psi} \|x - (\psi(\phi(x)))\|^2, x \in U$ // Use all unlabeled samples to train an autoencoder.
- 3: $\{\mu_1, \dots, \mu_b\} \leftarrow \arg \min_{\mu_i} \sum_{i=1}^k \sum_{x \in U_0^n} \|\phi(x) - \mu_i\|^2$ // Conduct K-means (set $k = b$) on the representations $\phi(x)$ to get b cluster centers.
- 4: Samples closest to each cluster center are selected and labeled to construct the unlabeled pool U_1^{n-b} and initial training dataset L_1^b .
- 5: $\mathcal{M}_{\theta_1} \leftarrow L_1^b$ // Use initial training set to train the proposed dual-channel DL model.
- 6: **while** $t \leq T$ and $Acc_{test} < Acc_{tar}$ **do**
- 7: $t \leftarrow t + 1$
- 8: $U_{n-b \times t}^t, L_{b \times t}^t \leftarrow Q(U_{n-b \times (t-1)}^{t-1}, L_{b \times (t-1)}^{t-1}, \mathcal{M}_{\theta_{t-1}})$ // Use the designed query strategy to select b useful unlabeled samples and label them.
- 9: $\mathcal{M}_{\theta_t} \leftarrow L_{b \times t}^t$ // Use current labeled training set with the learning algorithm to train the data-driven model.
- 10: $Acc_{test} \xleftarrow{D^T} \mathcal{M}_{\theta_t}$ // Validating the trained data-driven model on the test set.
- 11: **end while**
- 12: **return** $\mathcal{M}_{\theta_t}$

3) are interconnected to simulate different ITSCs (ITSC-#1-#3, #1-#2, and #2-#3). Various resistances (40/20/10/0.1/0 Ω) are connected to simulate winding SC faults in different degrees where lower resistance values correspond to more severe faults [26]. The wiring diagram is presented in Fig. 7, which shows the end-to-end open-circuit connection and records input and output signals, followed by applying the Fourier transform to obtain FRA gain and phase through formulas (1) and (2). Several typical FRA data, including gain and phase, are shown in Fig. 8. In contrast to data acquired

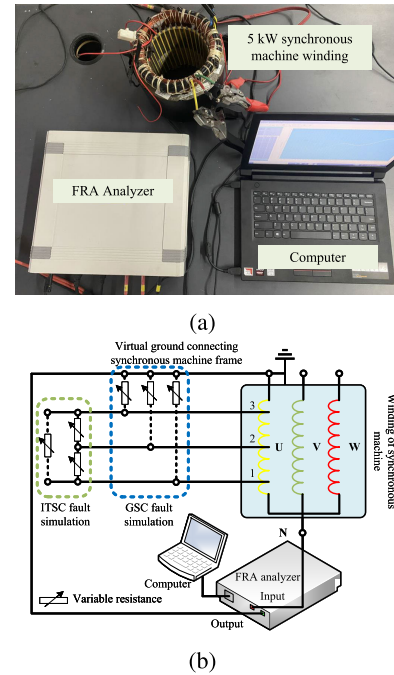


Fig. 7. Measurement experimental diagram. (a) Actual wiring diagram. (b) Measurement wiring diagram.

from normal windings, the FRA data associated with GSCs exhibit primary variations in the low-frequency band, whereas IDSCs manifest their dominant FRA data alterations in the midfrequency band. These observed trends are consistent with findings reported in [24], [27], and [32]. The experiments are repeated to create an FRA dataset consisting of seven classes, with 6021 samples allocated to the training set and 155 samples to the testing set. Details of the constructed dataset are provided in Table II.

With regard to the experimental setup, it should be noted that this study employs connecting resistances ranging from 0.1 to 40 Ω to simulate varying degrees of synchronous machine winding SC faults, solely to validate the proposed method. It is acknowledged that the specific resistance values should be adapted based on the characteristics of individual machines and the nature of actual SC fault conditions. Because prior studies [29], [32] have utilized resistances ranging from 1 to 100 Ω , this study, while informed by these studies,

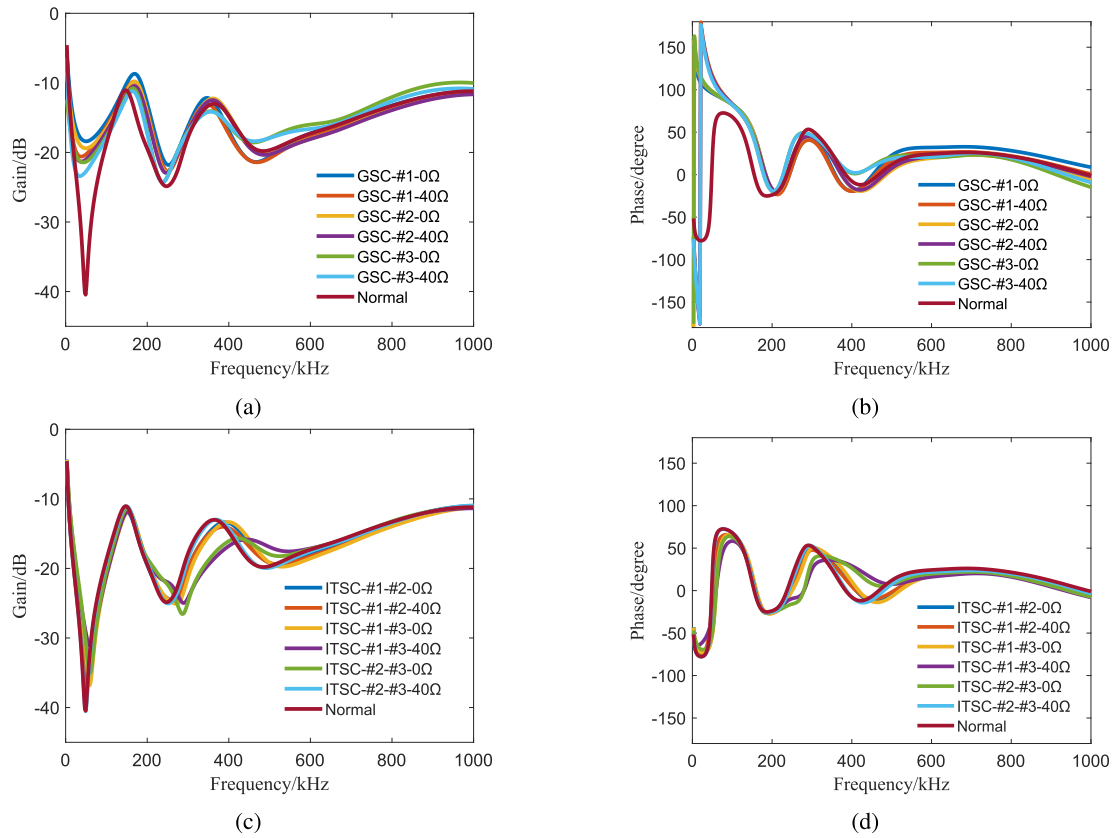


Fig. 8. Several FRA data of typical winding SC faults. (a) FRA gain of GSC. (b) FRA phase of GSC. (c) FRA gain of ITSC. (d) FRA phase of ITSC.

TABLE I
NAMEPLATE VALUE OF THE SYNCHRONOUS MACHINE
AND THE FRA ANALYZER

Characteristics	Parameter value
Rated power	5kVA
Rated voltage	380V
Frequency	50Hz
Pole pairs	1
Number of slots	36
Rated speed	1500rpm
FRA analyzer	
Model	TDT6U
Manufacturer	Beijing Shengtai Real-Time
Output	25 Vpp (maximum, adjustable)
Output impedance	1 MΩ / 50Ω (optional)
Frequency sweep range	1 Hz to 2 MHz
Frequency accuracy	99.995%
Dynamic range	-120 dB ~ 20 dB
Gain accuracy	± 0.5 dB
Test speed	no more than 1 minute

employs the 0.1–40 Ω. However, the main focus of this study remains the development of a winding SC fault diagnosis methodology, and comprehensive details regarding the specific experimental settings of FRA can be found in [10]. All codes utilized in this study are executed on the software and hardware configurations detailed in Table III. In addition, codes and hyperparameter settings can be found in the GitHub repository associated with this study.¹

¹(<https://github.com/cy1034429432/Intelligent-FRA-guided-Synchronous-Machine-Winding-SC-Faults-Diagnosis-Based-on-Active-Learning>)

IV. EXPERIMENT

A. Results of the Ablation Study

The model performance when using all samples as the training set is the upper bound for the model combined with AL [20]. Therefore, experiments are conducted utilizing the proposed dual-channel DL-, ML-based, and classical DL-based models. Except for the proposed model, other models concatenate FRA gain and phase images into one image as input, and the corresponding results are presented in Table IV. Unless otherwise specified, all reported accuracies (Tables IV–VI and Figs. 9–11) represent the average over 20 runs with different random seeds. Table IV shows that: 1) the ML model underperforms the DL model due to its limitations in handling high-dimensional image data; 2) increased model complexity does not necessarily improve performance given the limited dataset size; 3) the proposed model achieves comparable accuracy to state-of-the-art models [18], [26] while using significantly fewer parameters (1/500th of ResNet-18); and 4) utilizing dual-channel input (both FRA gain and phase) improves model performance compared to single-channel approaches (i.e., only gain or phase input). Table V shows the performance evaluation of the fault diagnosis model under varying module configurations and hyperparameter settings after ten training epochs. Different module configurations, including activation functions and neural network architectures, can significantly influence the convergence behavior of the model. Moreover, excessively large learning rates may induce convergence

TABLE II
FRA DATASET OF SYNCHRONOUS MACHINE WINDING

Fault type	Degrees	Training set	Test set
GSC-#1	GSC-#1-0Ω	196	5
	GSC-#1-0.1Ω	168	5
	GSC-#1-10Ω	195	5
	GSC-#1-20Ω	195	5
	GSC-#1-40Ω	195	5
GSC-#2	GSC-#2-0Ω	195	5
	GSC-#2-0.1Ω	195	5
	GSC-#2-10Ω	195	5
	GSC-#2-20Ω	197	5
	GSC-#2-40Ω	195	5
GSC-#3	GSC-#3-0Ω	195	5
	GSC-#3-0.1Ω	195	5
	GSC-#3-10Ω	195	5
	GSC-#3-20Ω	196	5
	GSC-#3-40Ω	195	5
ITSC-#1-#2	ITSC-#1-#2-0Ω	195	5
	ITSC-#1-#2-0.1Ω	195	5
	ITSC-#1-#2-10Ω	197	5
	ITSC-#1-#2-20Ω	195	5
	ITSC-#1-#2-40Ω	195	5
ITSC-#1-#3	ITSC-#1-#3-0Ω	195	5
	ITSC-#1-#3-0.1Ω	195	5
	ITSC-#1-#3-10Ω	194	5
	ITSC-#1-#3-20Ω	195	5
	ITSC-#1-#3-40Ω	195	5
ITSC-#2-#3	ITSC-#2-#3-0Ω	195	5
	ITSC-#2-#3-0.1Ω	195	5
	ITSC-#2-#3-10Ω	195	5
	ITSC-#2-#3-20Ω	197	5
	ITSC-#2-#3-40Ω	195	5
Normal		195	5
Total		6021	155

TABLE III
COMPUTER HARDWARE AND SOFTWARE CONFIGURATIONS

Device or software	Nameplate value or configuration
CPU	i5-12400F
GPU	RTX 3060
RAM	16G
Software	Python
Package	Pytorch, torchcam, sklearn

oscillations, whereas excessively small learning rates may impede convergence within the limited ten epochs.

The relationship between the number of labeled samples and the test set accuracy of the proposed dual-channel DL model under different query strategies is depicted in Fig. 9. Fig. 9 shows: 1) random sampling requires significantly more data to achieve comparable accuracy; 2) initializing the training set using an autoencoder and K -means clustering ensures sample diversity, leading to high initial accuracy. Subsequently, even with random sampling, satisfactory performance is achieved with fewer samples than with purely random sampling throughout the training process; 3) the two-stage strategy consistently outperforms single-stage strategies, achieving higher accuracy with fewer samples and exhibiting greater robustness to overfitting.

The encoder and the dual-channel model share the same architecture, enabling intuitive fine-tuning, analogous to pretraining and fine-tuning in large language models. However, as shown in Fig. 10, fine-tuning requires more labeled data to achieve comparable accuracy to training from scratch. This is

TABLE IV
COMPARISON OF DIFFERENT BACKBONE MODELS

Model	Test set accuracy	Model parameters
SVM	78.06%	\
KNN	71.61%	
Decision tree	75.48%	
AlexNet	86.26%	61107847
Vgg-11	88.38%	132864043
Vgg-16	96.13%	138358251
Vgg-19	83.87%	143667947
ResNet-18 (backbone model in previous work [18])	99.35%	11696519
ResNet-34	99.35%	21804679
ResNet-50	97.41%	25564039
iCaRLNet (backbone model in previous work [26])	99.35%	464089
AlexNet (only gain input)	80.64%	\
Vgg-11 (only gain input)	83.87%	
ResNet-18 (only gain input)	96.77%	
AlexNet (only phase input)	81.96%	\
Vgg-11 (only phase input)	84.51%	
ResNet-18 (only phase input)	94.19%	
Proposed model	99.35%	20923

TABLE V
COMPARISON OF DIFFERENT MODULE CONFIGURATIONS AND HYPERPARAMETER SETTINGS

Model	Test set accuracy
Backbone model Without IN	86.35%
Replace LeakyReLU with ReLU	92.25%
Replace CNN with FC	86.65%
Proposed model (Learning rate=0.001)	96.12%
Proposed model (Learning rate=0.005)	99.35%
Proposed model (Learning rate=0.01)	70.96%

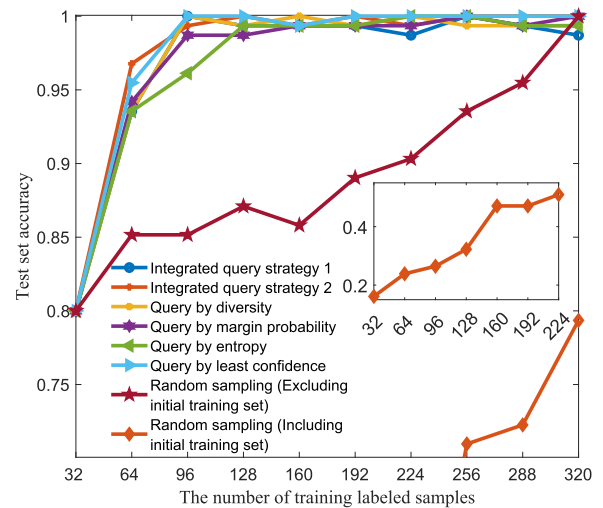


Fig. 9. Test set accuracy of training the proposed model from scratch under different numbers of samples selected by the mentioned query strategies. Including training the initial dataset means using representation learning to select samples and create the initial dataset, and the number of samples equals the value of batch size while excluding training the initial dataset means using random sampling to create the initial dataset.

attributed to the limitations of fine-tuning with the relatively large learning rate (0.005) required for rapid convergence within the constrained ten training epochs. Therefore, training the dual-channel model from scratch is recommended.

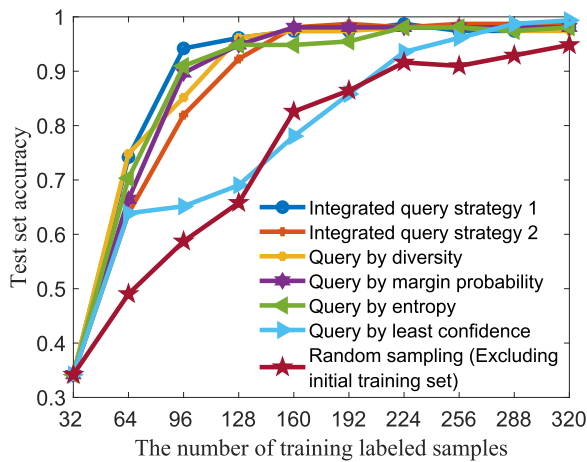


Fig. 10. Test set accuracy of fine-tuning the encoder under different numbers of samples selected by the mentioned query strategies.

Epoch and batch size significantly impact both training time and model performance, as shown in Fig. 11. From Fig. 11, it can be seen that: 1) results indicate a positive correlation between these hyperparameters and accuracy, but too large batch sizes do not improve the performance obviously because increasing batch size does not proportionally increase model updates, unlike increasing epochs; and 2) training time scales linearly with both epoch and batch size. Given the study's relatively large learning rate (0.005), a small epoch should be chosen to mitigate oscillations during training. Therefore, this study recommends that the epoch and batch size be set to 10 and 32, respectively.

B. Interpretability of the Trained Model

Researchers have been impressed by the effectiveness of DL methods while simultaneously raising concerns regarding their opaque nature. Understanding how a trained model works is crucial in determining whether it accurately captures feature information between different winding faults. This study applies t-distributed stochastic neighbor embedding (t-SNE) to the output of the trained dual-channel DL model on the test set, reducing its output to two dimensions. The visualization result is presented in Fig. 12. Fig. 12 shows that the trained model does learn to distinguish between the different winding faults using FRA gain and phase data, indicating it has captured essential features for classification.

Furthermore, to elucidate the decision-making criteria employed by the trained model for different winding fault types, this study applies Smooth Grad-CAM++ to visualize the sensitivity of the CNN module to specific image regions through gradient analysis, and its visualization results are presented in Fig. 13. Figs. 8 and 13 show that the trained dual-channel DL model does capture the differences between different classes. The visualization results reveal that the key to accurate FRA-guided synchronous machine winding fault diagnosis lies within the input of low- and mid-frequency bands. This observation aligns with the findings reported in [18].

C. Comparison Results of the Proposed Method With Previous Works

This study compares the performance of the proposed method with previous works [16], [18], [26]. Although these works differ in model inputs and datasets, they all revolve around the synchronous machine winding fault diagnosis guided by FRA. While it must be acknowledged that there exists an unfairness in conducting these comparisons, it is essential to highlight that the training samples, time, and epochs are significantly smaller than those needed in the previous works. When the model's accuracy on the test set reaches 99.35%, remarkable advantages can be observed compared to the previous methods. Specifically, the proposed method requires only 1/500th of the training time and 1/20th of the training samples compared to the previous DL-based model. Similarly, compared to the previous ML-based model, the proposed method demonstrates superior performance regarding reduced training time and the number of required labeled samples. Moreover, it should be noted that, although prior work employing the isolation forest [16] may yield comparable results, it cannot detect winding fault types. Compared to the isolation forest, the proposed models not only have additional functions (the ability to detect the types and locations of winding faults) but also require fewer training samples.

Remarkably, the proposed approach effectively reduces the burden on data annotation and the computing power required, all while satisfying the diagnostic model performance requirement ($\geq 95\%$).

V. DISCUSSION AND LIMITATION

Currently, numerous works have been done on synchronous machine winding fault diagnosis based on the data-driven model, but most works focus on the diagnostic model's performance without considering the cost of data annotation and computational power consumption. In real-world fault detection, applying data-driven models presents unique challenges, including a scarcity of experimental data and a high proportion of unlabeled samples. This study proposes a solution by efficiently selecting representative unlabeled samples, subsequently labeled by experienced experts. AL reduces the burden on data annotation and accelerates the training process.

In addition, this study uses image data of FRA gain and phase as input, but differences in winding types and turns among various machines may result in distinct FRA data. Although the trained model is not directly applied to different machines, its data-efficient nature allows rapid and cost-effective adaptation to various machines. To be specific, owing to the data efficiency of the proposed method, researchers augment the dataset with a small quantity of labeled data (approximately ten samples) from either new or identical machine types, compared with the thousands of labeled samples reported in previous works [18], [26]. Consequently, even if a new machine does not contribute fault data, similar samples can be obtained from other machines of the same type. On the other hand, machine manufacturers possess a large number of unlabeled data, so the proposed

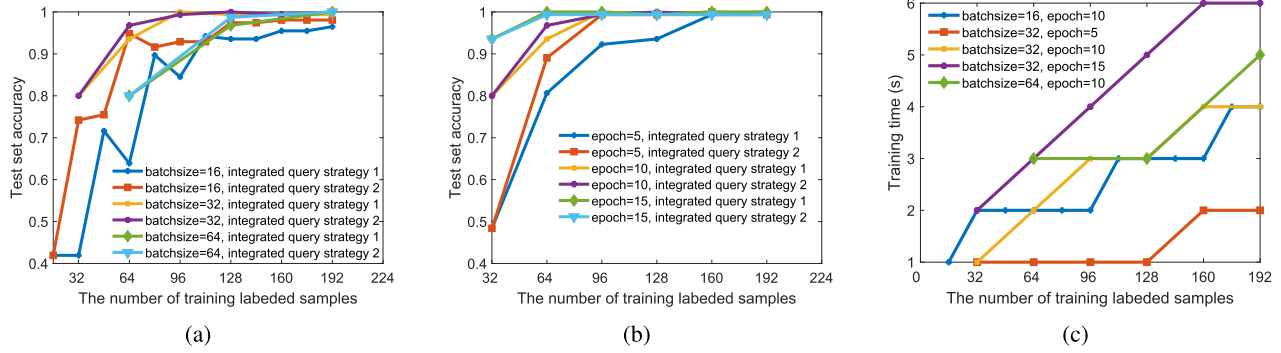


Fig. 11. Test set accuracy and training time under different training settings. (a) Test set accuracy under different batch sizes. (b) Test set accuracy under different epochs. (c) Training time under different settings.

TABLE VI
COMPARISON RESULTS WITH PREVIOUS WORKS

Method	Training time	Training set samples	Test set samples	Training epoch	Test set accuracy	Whether can detect fault types
Isolation forest [16]	6.32s	320	44	30	100%	×
ResNet-18 [18]	29min30s	1808	270	100	99.63%	✓
iCaRLNet with life long learning (one class per task) [26]	27min59s	1808	270	70	92.59%	✓
Integrated query strategy 1 (epoch=10 batchsize=32)	2s	64	155	10	93.55%	✓
Integrated query strategy 1 (epoch=10 batchsize=32)	3s	96	155	10	99.35%	✓
Integrated query strategy 2 (epoch=10 batchsize=32)	2s	64	155	10	96.7%	✓
Integrated query strategy 2 (epoch=10 batchsize=32)	3s	96	155	10	99.35%	✓

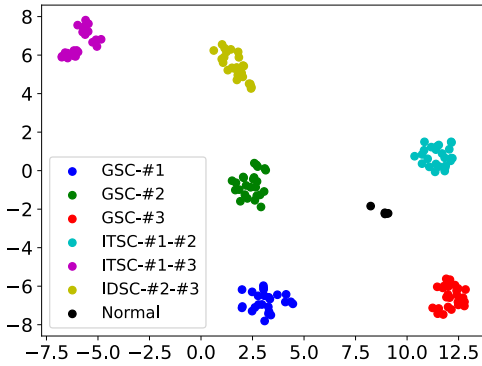


Fig. 12. Visualization results of the test set based on the trained dual-channel DL model and t-SNE.

method entails selecting representative samples for expert labeling. Subsequently, these newly annotated data augment the dataset for model training, enabling the development of fault diagnosis models tailored to various machines through the application of transfer learning [33] and lifelong learning [26]. Based on the proposed method, each machine is expected to have a customized fault diagnosis model. Currently, FRA is a prevalent technique for synchronous machine inspection during scheduled shutdown maintenance. Therefore, the customized fault diagnosis model offers a means to streamline this process and support more objective decision-making.

While the proposed method demonstrates advantages in data annotation and accelerates model training, it still has several limitations.

- 1) Due to experimental constraints, the dataset used only includes data about manually simulated winding SC faults on a 5-kVA synchronous machine. As a result, the model derived from this dataset only applies to the experimental machine. It is advisable to incorporate data from a broader range of machines to improve the model's generalizability, even including larger industrial machines and different machine types. Furthermore, the experimental data utilized in this study are derived from manually simulated winding faults rather than real-world fault data. Future research should incorporate data acquired from actual fault events to enhance the practical applicability and robustness of the proposed model.
- 2) The essence of FRA is to compare normal and faulty FRA data, and the FRA-guided fault diagnosis models are still suitable for offline detection. Due to some differences in the definition of fault severity among different machines, it is still necessary to train models from scratch based on corresponding data for different machines.
- 3) This study has not applied the proposed method to real-world fault detection scenarios. Future research should prioritize the integration of the trained model into embedded systems, potentially leveraging edge computing platforms. In the context of edge computing, critical considerations include computational resources and storage capacity of hardware, as well as the feasibility of model compression techniques [34], [35].

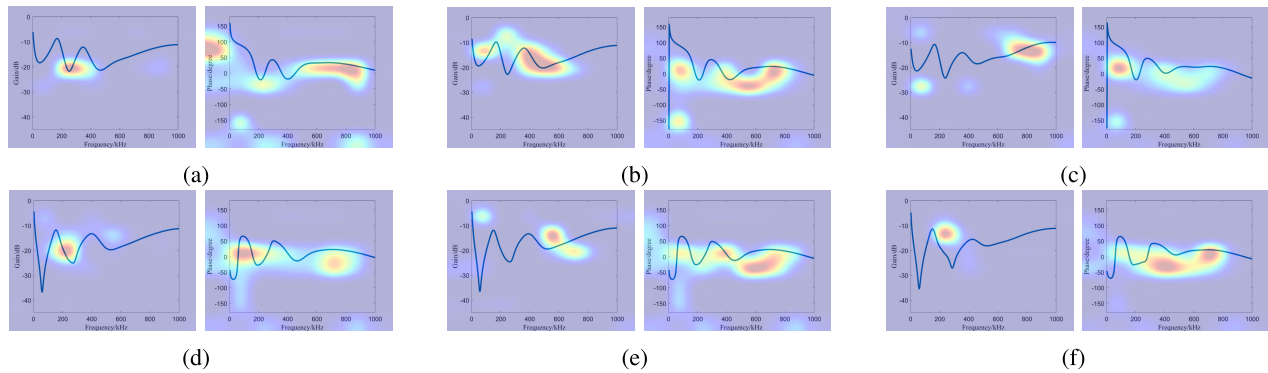


Fig. 13. Visualization results based on the proposed model and Smooth Grad-CAM++. (a) GSC-#1. (b) GSC-#2. (c) GSC-#3. (d) ITSC-#1-#3. (e) ITSC-#1-#2. (f) ITSC-#2-#3. The redder the color of the area, the greater the impact on the trained model, which means the trained model makes decisions based on data from darker-colored areas.

VI. CONCLUSION

This study proposes a data-efficient synchronous machine winding SC faults diagnosis method based on FRA and AL. According to the experimental and comparative results, the following conclusions are obtained.

- 1) Efficient model training requires both representative sample selection for the training set and a lightweight backbone model tailored to dataset size, and performance gains may be more effectively achieved through additional input signals rather than increased model parameters. Therefore, classical models are often unsuitable for electrical equipment fault diagnosis, hindering their practical application.
- 2) The two-stage query strategy mitigates the limitations of both uncertainty- and diversity-based methods by avoiding outlier selection and ensuring diversity, so the proposed method substantially reduces both data annotation costs and computational demands. Furthermore, unsupervised representation learning before AL benefits the initial training set, aligning with the prevalence of unlabeled samples in real-world scenarios.

ACKNOWLEDGMENT

The authors would like to thank Dr. Zhongtuo Shi at the School of Electrical and Electronic Engineering, Huazhong University of Science and Technology (HUST), who gave helpful suggestions about applying AL to winding SC faults diagnosis.

REFERENCES

- [1] K. N. Gyftakis and A. J. Marques-Cardoso, "Reliable detection of very low severity level stator inter-turn faults in induction motors," in *Proc. 45th Annu. Conf. IEEE Ind. Electron. Soc.*, Oct. 2019, pp. 1290–1295.
- [2] S. Zhang, S. Li, L. He, J. A. Restrepo, and T. G. Habetler, "A high-frequency rotating flux injection based rotor thermal monitoring scheme for direct-torque-controlled interior permanent magnet synchronous machines," in *Proc. IEEE Int. Electric Mach. Drives Conf. (IEMDC)*, May 2017, pp. 1–6.
- [3] M. S. Rafiq et al., "A simple method for identifying mass unbalance using vibration measurement in permanent magnet synchronous motors," *IEEE Trans. Ind. Electron.*, vol. 69, no. 6, pp. 6441–6444, Jun. 2022.
- [4] I. Zamudio-Ramirez, R. A. Osorio-Rios, J. A. Antonino-Daviu, H. Razik, and R. D. J. Romero-Troncoso, "Magnetic flux analysis for the condition monitoring of electric machines: A review," *IEEE Trans. Ind. Informat.*, vol. 18, no. 5, pp. 2895–2908, May 2022.
- [5] G. Niu, X. Dong, and Y. Chen, "Motor fault diagnostics (Don't short) based on current signatures: A review," *IEEE Trans. Instrum. Meas.*, vol. 72, pp. 1–19, 2023.
- [6] T. Li et al., "Simulation study on interturn short circuit of rotor windings in generator by RSO method," in *Proc. IEEE Int. Conf. High Voltage Eng. Appl. (ICHVE)*, Sep. 2018, pp. 1–4.
- [7] G.-A. Capolino and A. Cavagnino, "New trends in electrical machines technology—Part II," *IEEE Trans. Ind. Electron.*, vol. 61, no. 9, pp. 4931–4936, Sep. 2014.
- [8] C. Martin, J. Guerrero, P. Gomez-Mourello, and C. Platero, "Ground faults location in poles of synchronous machines through frequency response analysis," *IEEE Trans. Ind. Appl.*, vol. 58, no. 1, pp. 113–122, Jan./Feb. 2022.
- [9] H. Mayora, R. Alvarez, G. Bossio, and E. Calo, "Condition assessment of rotating electrical machines using SFRA—A survey," in *Proc. IEEE Electr. Insul. Conf. (EIC)*, Jun. 2021, pp. 26–29.
- [10] Y. Yu, Z. Zhao, Y. Chen, H. Wu, C. Tang, and W. Gu, "Evaluation of the applicability of IFRA for short circuit fault detection of stator windings in synchronous machines," *IEEE Trans. Instrum. Meas.*, vol. 71, pp. 1–12, 2022.
- [11] A. Mugarra, H. Mayora, J. M. Guerrero, and C. A. Platero, "Frequency response analysis (FRA) fault diagram assessment method," *IEEE Trans. Ind. Appl.*, vol. 58, no. 1, pp. 336–344, Jan. 2022.
- [12] L. Wang, C. Ngai-Man Ho, F. Canales, and J. Jatskevich, "High-frequency modeling of the long-cable-fed induction motor drive system using TLM approach for predicting overvoltage transients," *IEEE Trans. Power Electron.*, vol. 25, no. 10, pp. 2653–2664, Oct. 2010.
- [13] J. Ni, Z. Zhao, S. Tan, Y. Chen, C. Yao, and C. Tang, "The actual measurement and analysis of transformer winding deformation fault degrees by FRA using mathematical indicators," *Electr. Power Syst. Res.*, vol. 184, Jul. 2020, Art. no. 106324.
- [14] M. H. Samimi and S. Tenbohlen, "FRA interpretation using numerical indices: State-of-the-art," *Int. J. Electr. Power Energy Syst.*, vol. 89, pp. 115–125, Jul. 2017.
- [15] Y. Chen, Z. Zhao, J. Liu, S. Tan, and C. Liu, "Application of generative AI-based data augmentation technique in transformer winding deformation fault diagnosis," *Eng. Failure Anal.*, vol. 159, May 2024, Art. no. 108115.
- [16] Y. Chen, Z. Zhao, H. Wu, X. Chen, Q. Xiao, and Y. Yu, "Fault anomaly detection of synchronous machine winding based on isolation forest and impulse frequency response analysis," *Measurement*, vol. 188, Jan. 2022, Art. no. 110531.
- [17] L. Zhou, T. Lin, X. Zhou, S. Gao, Z. Wu, and C. Zhang, "Detection of winding faults using image features and binary tree support vector machine for autotransformer," *IEEE Trans. Transport. Electrification*, vol. 6, no. 2, pp. 625–634, Jun. 2020.
- [18] Y. Chen, Z. Zhao, Y. Yu, W. Wang, and C. Tang, "Understanding IFRA for detecting synchronous machine winding short circuit faults based on image classification and smooth grad-CAM++," *IEEE Sensors J.*, vol. 23, no. 3, pp. 2422–2432, Feb. 2023.
- [19] B. Settles, "Active learning literature survey," *Comput. Sci., Univ. Wisconsin–Madison, Madison, WI, USA, Tech. Rep. 1648*, 2009.

- [20] Z. Shi, W. Yao, Y. Tang, X. Ai, J. Wen, and S. Cheng, "Intelligent power system stability assessment and dominant instability mode identification using integrated active deep learning," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 35, no. 7, pp. 9970–9984, Jul. 2024.
- [21] P. Ren et al., "A survey of deep active learning," *ACM Comput. Surv.*, vol. 54, no. 9, p. 180, 2021.
- [22] Y. Chen, B. Qin, X. Liu, W. Wang, and Y. Liao, "A data-efficient surrogate modeling method for a cyclotron-based proton therapy beamline based on active learning," *Int. J. Mod. Phys. A*, vol. 40, no. 1, Jan. 2025, Art. no. 2450160.
- [23] Z. Shi, W. Yao, Y. Tang, X. Ai, J. Wen, and S. Cheng, "Bidirectional active transfer learning for adaptive power system stability assessment and dominant instability mode identification," *IEEE Trans. Power Syst.*, vol. 38, no. 6, pp. 5128–5142, Nov. 2023.
- [24] Z. Zhao, Y. Chen, Y. Yu, M. Han, C. Tang, and C. Yao, "Equivalent broadband electrical circuit of synchronous machine winding for frequency response analysis based on gray box model," *IEEE Trans. Energy Convers.*, vol. 36, no. 4, pp. 3512–3521, Dec. 2021.
- [25] H. Badihi, Y. M. Zhang, B. Jiang, P. Pillay, and S. Rakheja, "A comprehensive review on signal-based and model-based condition monitoring of wind turbines: Fault diagnosis and lifetime prognosis," *Proc. IEEE*, vol. 110, no. 6, pp. 754–806, Jun. 2022.
- [26] Y. Chen, Z. Zhao, Y. Yu, Y. Guo, and C. Tang, "Improved interpretation of impulse frequency response analysis for synchronous machine using life long learning based on iCaRL," *IEEE Trans. Instrum. Meas.*, vol. 72, pp. 1–10, 2023.
- [27] A. Mugarra, H. Mayora, J. M. Guerrero, and C. A. Platero, "Validity of frequency response analysis (FRA) for diagnosing large salient poles of synchronous machines," *IEEE Trans. Ind. Appl.*, vol. 58, no. 1, pp. 226–234, Jan./Feb. 2022.
- [28] F. R. Blázquez, C. A. Platero, E. Rebollo, and F. Blázquez, "Novel rotor ground-fault detection algorithm for synchronous machines with static excitation based on third-harmonic voltage-phaser comparison," *IEEE Trans. Ind. Electron.*, vol. 63, no. 4, pp. 2548–2558, Apr. 2016.
- [29] F. R. Blázquez, Carlos. A. Platero, E. Rebollo, and F. Blázquez, "Evaluation of the applicability of FRA for inter-turn fault detection in stator windings," in *Proc. 9th IEEE Int. Symp. Diagnostics Electr. Mach., Power Electron. Drives (SDEMPED)*, Aug. 2013, pp. 177–182.
- [30] F. R. Blázquez, C. A. Platero, E. Rebollo, and F. Blázquez, "Field-winding fault detection in synchronous machines with static excitation through frequency response analysis," *Int. J. Electr. Power Energy Syst.*, vol. 73, pp. 229–239, Dec. 2015.
- [31] X. Liu, W. Miao, Q. Xu, L. Cao, C. Liu, and P. W. T. Pong, "Inter-turn short-circuit fault detection approach for permanent magnet synchronous machines through stray magnetic field sensing," *IEEE Sensors J.*, vol. 19, no. 18, pp. 7884–7895, Sep. 2019.
- [32] S. Uhrig, F. Öttl, R. Hinterholzer, and N. Augener, "Reliable diagnostics on rotating machines using FRA," in *Proc. Int. Conf. Diagnostics Electr. Eng. (Diagnostics)*, Sep. 2020, pp. 1–6.
- [33] X. Chen, Z. Zhao, F. Guo, S. Tan, and J. Wang, "Diagnosis method of transformer winding mechanical deformation fault based on sliding correlation of FRA and series transfer learning," *Electr. Power Syst. Res.*, vol. 229, Apr. 2024, Art. no. 110173.
- [34] J. Lu, K. An, X. Wang, J. Song, F. Xie, and S. Lu, "Compressed channel-based edge computing for online motor fault diagnosis with privacy protection," *IEEE Trans. Instrum. Meas.*, vol. 72, 2023, Art. no. 6505112.
- [35] C. He et al., "Real-time fault diagnosis of motor bearing via improved cyclostationary analysis implemented onto edge computing system," *IEEE Trans. Instrum. Meas.*, vol. 72, 2023, Art. no. 3524011.



Yu Chen was born in Wenzhou, Zhejiang, China, in 2000. He received the B.S. degree from the College of Engineering and Technology, Southwest University, Chongqing, China, in 2022. He is currently pursuing the master's degree with the School of Electrical and Electronic Engineering, Huazhong University of Science and Technology, Wuhan, Hubei, China.

At the same time, he is a part-time Research Assistant at the the Department of Building Environment and Energy Engineering, The Hong Kong Polytechnic University. His research areas include condition monitoring, fault diagnosing for power equipment, and the application of artificial intelligence.



Zhongyong Zhao (Member, IEEE) was born in Guangyuan, Sichuan, China. He received the B.S. and Ph.D. degrees in electrical engineering from Chongqing University, Chongqing, China, in 2011 and 2017, respectively.

His areas of research include condition monitoring and fault diagnosing for high-voltage apparatus and artificial intelligence.

Dr. Zhao received a Scholarship from China Scholarship Council to enable him to attend a joint-training Ph.D. Program at Curtin University, Perth, WA, Australia, from 2015 to 2016.



Jiangnan Liu was born in Ganzhou, Jiangxi, China. He received the B.S. degree from the College of Engineering and Technology, Southwest University, Chongqing, China, in 2020, and the master's degree from the Department of Electrical Engineering, Chongqing University, Chongqing, in 2023, with a fault diagnosing for HV apparatus and pulsed power technology.

He currently works with Guangzhou Power Supply Bureau, Guangdong Power Grid Company Ltd., Guangdong, China.



Wei Wang was born in Shaoyang, Hunan, China, in 1998. He received the B.S. degree from the School of Electrical and Information Engineering, Changsha University of Science and Technology, Changsha, China, in 2019. He is currently pursuing the Ph.D. degree with the School of Electrical and Electronic Engineering, Huazhong University of Science and Technology, Wuhan, Hubei, China.

His research areas include proton therapy, medical physics, and the application of deep learning.



Chenguo Yao (Member, IEEE) was born in Nanchong, Sichuan, China. He received the B.S., M.S., and Ph.D. degrees in electrical engineering from Chongqing University, Chongqing, China, in 1997, 2000, and 2003, respectively.

He was a Professor at the College of Electrical Engineering, Chongqing University, in 2007. His research interests include pulse power technology and its application in biomedical engineering, online monitoring of insulation conditions, and insulation fault diagnosis for HV apparatus.