MITSUBISHI ELECTRIC RESEARCH LABORATORIES
https://www.merl.com

Simulation-to-Reality Domain Adaptation for Motor Fault Detection

Ji, Dai-Yan; Wang, Bingnan; Inoue, Hiroshi; Kanemaru, Makoto

TR2025-126 August 26, 2025

Abstract

In this paper, we investigate a simulation-to-reality domain adaptation approach for detecting motor faults, aiming to address the data scarcity problem in data-driven fault detection. A physics-based fault model is developed to generate simulation data under various fault conditions. A DQ transformation and feature extraction step is then performed for both simulation and real data, before domain adaptation is applied to align the simulation data with the limited real measurement data. Machine learning models can then be trained on the adapted data to make predictions. We demonstrate the effectiveness of the proposed method on eccentricity fault level prediction of an induction motor using stator current signal. Results demonstrate superior prediction performance compared to baseline model with only real data, with significant error reduction under both no-load and on-load conditions. This approach offers a promising and practical solution for motor fault detection in scenarios where obtaining comprehensive real fault data is challenging, as it leverages simulated data to enhance model performance with limited real-world measurements.

IEEE International Symposium on Diagnostics for Electric Machines, Power Electronics and Drives (SDEMPED) 2025

Mitsubishi Electric Research Laboratories, Inc.
201 Broadway, Cambridge, Massachusetts 02139

Simulation-to-Reality Domain Adaptation for Motor Fault Detection

Dai-Yan Ji
Mechanical Engineering
University of Maryland
College Park, MD 20742, USA
jidn@umd.edu

Bingnan Wang
Mitsubishi Electric Research Laboratories
Cambridge, MA 02139, USA
bwang@merl.com

Hiroshi Inoue
Advanced Technology R&D Center
Mitsubishi Electric Corporation
Amagasaki, Hyogo 661-8661 Japan

Makoto Kanemaru
Advanced Technology R&D Center
Mitsubishi Electric Corporation
Amagasaki, Hyogo 661-8661 Japan

***Abstract*—In this paper, we investigate a simulation-to-reality domain adaptation approach for detecting motor faults, aiming to address the data scarcity problem in data-driven fault detection. A physics-based fault model is developed to generate simulation data under various fault conditions. A DQ transformation and feature extraction step is then performed for both simulation and real data, before domain adaptation is applied to align the simulation data with the limited real measurement data. Machine learning models can then be trained on the adapted data to make predictions. We demonstrate the effectiveness of the proposed method on eccentricity fault level prediction of an induction motor using stator current signal. Results demonstrate superior prediction performance compared to baseline model with only real data, with significant error reduction under both no-load and on-load conditions. This approach offers a promising and practical solution for motor fault detection in scenarios where obtaining comprehensive real fault data is challenging, as it leverages simulated data to enhance model performance with limited real-world measurements.**

***Index Terms*—Simulation-to-Reality (Sim2Real), domain adaptation, eccentricity fault detection, induction motors**

I. Introduction

Motor fault detection and predictive maintenance are important in protecting the assets in many industries. A lot of fault conditions can occurs during the lifetime of an electric motor. In particular, eccentricity is a significant indicator of mechanical faults [1], [2] in electric machines. It occurs when the stator and rotor are not concentric. This common fault can significantly impact motor performance and reliability. Traditionally, motor fault detection typically depends on sensing modalities such as vibration and acoustic emission [3]. However, these measurements are often unreliable due to sensor mounting locations and interference from nearby machinery, making it difficult to accurately detect faults or quantify their severity. Conversely, motor current signature analysis (MCSA) offers several advantages over vibration and acoustic sensing methods, including simple implementation and low cost [4]. MCSA relies on measured motor current data to detect motor faults [5]–[7], eliminating the need for additional sensors. Physics-based techniques [8]–[10] establish models to describe motor faults, identifying detailed fault signatures and specifying requirements for experimental measurements and data analysis to extract appropriate signals. However, there is always some discrepancies between signals generated by physical models and real measurement data, making it challenging to identify and quantify motor faults based on model only.

On the other hand, data-driven methods, especially machine learning methods [11], [12] rely on experimental data to train models that can make decisions regarding motor fault conditions. However, in practical applications and operational environments, particularly for motor eccentricity fault analysis, fault detection using data-driven techniques presents a significant challenge. The main issue is the lack of sufficient data to represent both good and faulty conditions (i.e., different eccentricity levels) because induction motors typically operate normally, making it difficult to obtain faulty data.

To address these challenges, we propose to apply simulation-to-reality (Sim-to-Real, Sim2Real) Domain Adaptation (DA) that employs DA algorithms [13] to transfer knowledge learned in a simulated environment to real-world applications. Sim2Real DA [14], [15] offers a number of advantages compared with other machine learning based methods, including reduced development costs and time (e.g., difficult to install experimental equipment and collect sufficient real data) and enhanced reliability and robustness during testing (e.g., generating simulated data with different eccentricity levels as source data to adapt and train a robust model).

In this paper, we apply Sim2Real DA to the eccentricity level prediction of an electric motor: first a physics model is developed to generate simulation data under faulty conditions; then a DQ transformation and feature extraction step is applied to both simulation and real data to reduce the dimension and

This work was done when D. Ji was with Mitsubishi Electric Research Laboratories as research intern. Corresponding Author: Bingnan Wang (bwang@merl.com).

extract fault-related information; domain adaptation is then applied to align the simulation data with the limited real measurement data; finally a regression model is trained on the adapted data to make predictions. Multiple numerical tests with two different domain adaptation techniques are performed to demonstrate the effectiveness of the method.

The remaining parts of this paper are organized as follows: Section III outlines the data acquisition and data analysis. Section IV details the experiment setup. Sections V–VI present results and conclude the paper.

II. Methodologies

The proposed methodology for detecting induction motor eccentricity faults is illustrated in Fig. 1. First, the simulated data is generated using a physics-based model, denoted as source domain data X_s. Then, data pre-processing and feature extraction is performed to obtain Z_s, in this paper through DQ transformation. On the other hand, limited real measurement data serve as target domain data X_t. With the same data pre-processing and feature extraction step, target domain data is obtained in feature space, denote as Z_t. Domain adaptation is performed to align the distribution of source domain data Z_s with target domain data Z_t, and the transformed source domain data is denoted as Z'_s after the DA process. During the training stage, a regression model is trained to minimize the prediction error compared against actual label of the training data. In the testing stage, real measurement data is fed into the trained model to make predictions on motor fault condition.

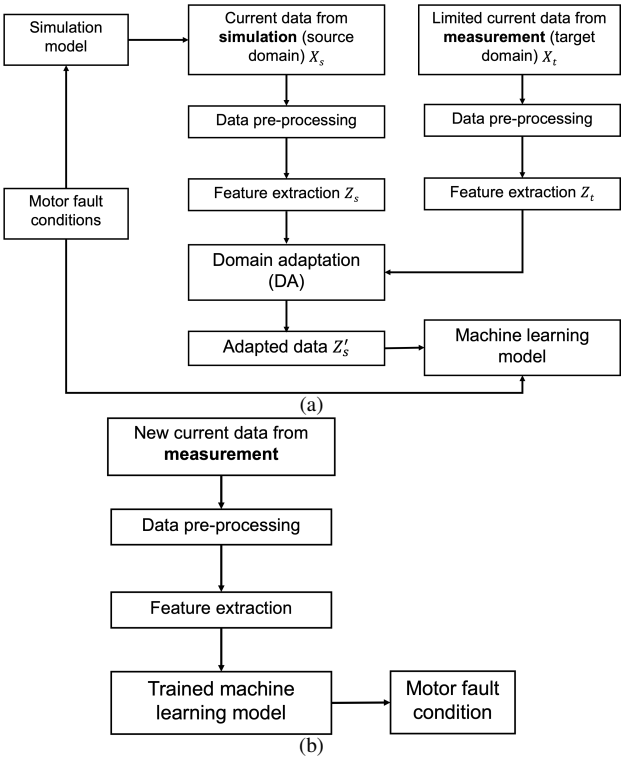

Fig. 1: Flow chart of the Sim2Real DA method for motor fault detection. (a) Training process. (b) Evaluation process of a trained model.

A. Physics-based Eccentricity Model

We develop a physics-based eccentricity model to generate simulation data under different eccentricity conditions, based on modified winding function method (MWFM) and coupled-circuit model. The key input parameters of this physics-based numerical model include: motor design parameters, supply voltage, load condition, and fault condition. These parameters are used to calculate the inductance terms between rotor and stator windings at each rotor position, updating the dynamic signals including stator current, speed, and torque throughout the motor operation. The eccentricity condition is described by period modulations to the air gap function $g(\phi, t)$:

$$g(\phi, t) = g_0 K_c - \delta_{\text{SE}} g_0 \cos(\phi) - \delta_{\text{DE}} g_0 \cos(\phi - \omega_r t) \quad (1)$$

with g_0 as the nominal air gap length, K_c as Carter's coefficient, and δ_{SE}, δ_{DE} as the static and dynamic eccentricity amplitudes, respectively.

The motor dynamics are represented through coupled circuit equations. Inductance terms and their derivatives, crucial for determining motor current and torque, are computed using MWFM, updating at each rotor position [5], [10], [16]. The detailed modeling process is further described in [17].

Dynamic simulations are subsequently conducted to obtain motor current signals under each condition. While the simulation does not exactly match the experimental data due to inevitable model simplifications, it accurately identifies key signal features attributable to eccentricity. The model describes the electromagnetic interactions within the motor under eccentric conditions, providing a robust framework for analyzing motor faults.

B. DQ Transformation & Feature Extraction

The DQ Transformation of three-phase currents is a mathematical method that converts three-phase alternating current (AC) signals into direct current (DC) form. This transformation converts the dynamic signals of a three-phase system into a rotating reference frame, where they appear as constant (DC-like) values under steady-state conditions on the rotating DQ axes, simplifying the control and analysis of these signals. The steps of this process are as follows:

1) Clarke Transformation (abc to $\alpha\beta$ Transformation): The Clarke transformation converts the three-phase signals i_a, i_b, and i_c into the $\alpha\beta$ reference frame i_α, i_β. The matrix form of the Clarke transformation is:

$$\begin{bmatrix} i_\alpha \\ i_\beta \end{bmatrix} = \frac{2}{3} \begin{bmatrix} 1 & -\frac{1}{2} & -\frac{1}{2} \\ 0 & \frac{\sqrt{3}}{2} & -\frac{\sqrt{3}}{2} \end{bmatrix} \begin{bmatrix} i_a \\ i_b \\ i_c \end{bmatrix} \quad (2)$$

2) Park Transformation ($\alpha\beta$ to DQ Transformation): Once the final rotation angle θ_{final} is calculated (based on the system's angular velocity and PLL adjustment), the Park transformation can be performed to convert the signals from the $\alpha\beta$ coordinate system to the DQ coordinate

system. The mathematical expression for this process is as follows:

$$\begin{bmatrix} i_d \\ i_q \end{bmatrix} = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} i_\alpha \\ i_\beta \end{bmatrix} \tag{3}$$

Phase angle θ is the instantaneous angular position of the rotating reference frame, and can be computed based on the system's angular velocity ωt, and dynamically adjusted using a phase-locked loop (PLL) to synchronize with the phase of the motor's rotating magnetic field.

3) Feature Extraction from i_d Signal: Since i_d signal is related to the magnetic flux, it provides critical information about the motor's magnetic characteristics to extract meaningful statistical features. We first calculate the mean of the i_d signal, and then subtract the mean from each signal to center the waveform around zero, highlighting deviations and variations and making waveform analysis more precise. Then, we extract eleven primary statistical features: mean, standard deviation, maximum, minimum, peak, root mean square (RMS), skewness, kurtosis, range, median, and interquartile range (IQR), as detailed in Table I.

C. Domain Adaptation

The goal of domain adaptation is to map the feature distributions from the source domain D_s to the target domain D_t, thereby reducing the distributional discrepancy between them. Features are extracted from both domains to create feature vectors X_s and X_t, which are subsequently aligned during the domain adaptation process.

Various feature alignment methods have been developed for domain adaptation, such as CORrelation ALignment (CORAL), Transfer Component Analysis (TCA), Subspace Alignment (SA), etc. In particular, CORAL is a popular DA method due to its simplicity in implementation and effectiveness in reducing domain mismatch [18].

More recently, optimal transport (OT) based methods have been applied to domain adaptation and show advantages in achieving precise alignment and handling more complex domain shifts and class imbalances [19]. In this paper, we implement two representative methods: CORAL and OT, for our domain adaptation process, and compare their performances for our motor fault detection problem.

1) CORelation ALignment (CORAL): To ensure that the alignment focuses solely on second-order statistics (i.e., the shape of the distribution), feature vectors X_s and X_t are first centered by subtracting theirs respective means. Then, the covariance matrices are computed with regularization λ for numerical stability:

$$\Sigma_s = \text{cov}(X_s - \mu_s) + \lambda I, \quad \Sigma_t = \text{cov}(X_t - \mu_t) + \lambda I \tag{4}$$

The transformation matrix A is computed to align the second-order statistics between domains:

$$A = \Sigma_s^{-\frac{1}{2}} \Sigma_t^{\frac{1}{2}} \tag{5}$$

Finally, the obtained matrix A is then used to transform the source domain features:

$$X_s^{'} = (X_s - \mu_s)A + \mu_t \tag{6}$$

2) Optimal Transport (OT): OT for domain adaptation aims to find a transportation plan that minimizes the cost of moving data points from the source domain to the target domain, effectively aligning their distributions. In particular, we formulate the problem as aregularized OT [19], where an entropy term is added to the classic transport cost function to improve computational efficiency and promote smoother solutions:

$$\gamma^* = \arg\min_{\gamma} \sum_{i,j} \gamma_{i,j} C(x_s^i, x_t^j) + \epsilon \sum_{i,j} \gamma_{i,j} \log(\gamma_{i,j}) \tag{7}$$

where γ represents the transport plan, $C(x_s^i, x_t^j)$ is the cost of transporting x_s^i to x_t^j, and ϵ is the regularization parameter that balances the transport cost and the entropy of the transport plan. The pairwise distances between the source and target domain data points are captured by the cost matrix C, typically computed using the squared Euclidean distance:

$$C_{i,j} = \|x_s^i - x_t^j\|^2 \tag{8}$$

While equation (7) defines the theoretical OT problem with entropy regularization, solving this problem directly is computationally intensive due to the high dimensionality of the data and the need for stability in finding the transport plan. To solve this problem, the Sinkhorn algorithm is designed to efficiently solve the regularized OT problem by iteratively adjusting the transport plan γ.

Using the cost matrix C and the regularization parameter ϵ, the Sinkhorn kernel K is computed as:

$$K = \exp\left(-\frac{C}{\epsilon}\right) \tag{9}$$

The Sinkhorn algorithm iteratively updates the dual variables u and v, which correspond to scaling factors for the source and target distributions, respectively. The update process is performed alternately: first, $u^{(k+1)}$ is updated based on the current $v^{(k)}$, and then $v^{(k+1)}$ is updated using the newly computed $u^{(k+1)}$. This alternating process continues until convergence. The update rules are given by:

$$u^{(k+1)} = \frac{\mathbf{r}}{K v^{(k)}}, \quad v^{(k+1)} = \frac{\mathbf{c}}{K^T u^{(k+1)}} \tag{10}$$

where $\mathbf{r} \in \mathbb{R}^{n_s}$ and $\mathbf{c} \in \mathbb{R}^{n_t}$ are uniform probability vectors. Specifically, $\mathbf{r} = [\mathbf{r}_1, \ldots, \mathbf{r}_i]^\top$ and $\mathbf{c} = [\mathbf{c}_1, \ldots, \mathbf{c}_j]^\top$, where each element $\mathbf{r}_i = 1/n_s$ for all i and $\mathbf{c}_j = 1/n_t$ for all j, with n_s and n_t denoting the number of source and target samples, respectively. After several iterations, the Sinkhorn algorithm converges, and the final transport plan γ^* is obtained. This solution can be expressed as:

$$\gamma^* = \text{diag}(u) K \text{diag}(v) \tag{11}$$

TABLE I: Features Extracted from i_d Signal

No.	Feature Name	Feature Description
1	$\mu = \frac{1}{n}\sum_{i=1}^{n} x_i$	Mean value of the i_d signal
2	$\sigma = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i - \mu)^2}$	Standard deviation of the i_d signal, measuring dispersion
3	Max $= \max(x_i)$	Maximum value of the i_d signal
4	Min $= \min(x_i)$	Minimum value of the i_d signal
5	Peak $= \max \lvert x_i \rvert$	Peak value, the maximum absolute deviation from zero in the i_d signal
6	RMS $= \sqrt{\frac{1}{n}\sum_{i=1}^{n} x_i^2}$	Root mean square (RMS) of the i_d signal, indicating the magnitude
7	Skewness $= \frac{\sum_{i=1}^{n}(x_i-\mu)^3}{n\sigma^3}$	Skewness of the i_d signal, indicating the asymmetry of the data distribution
8	Kurtosis $= \frac{\sum_{i=1}^{n}(x_i-\mu)^4}{n\sigma^4}$	Kurtosis of the i_d signal, measuring the heaviness of the tails
9	Range $= \max(x_i) - \min(x_i)$	Range of the i_d signal
10	Median $= Q_2$	Median value of the i_d signal, representing the central value
11	IQR $= Q_3 - Q_1$	Interquartile range (IQR), the difference between the third quartile Q_3 and the first quartile Q_1 of the i_d signal

Equation (11) represents the output of the Sinkhorn algorithm, which provides an efficient numerical solution to the regularized OT problem defined in equation (7). Using the computed transport plan γ^*, the source domain data X_s is transformed into the target domain space:

$$X_s^{'} = \gamma^* X_s \tag{12}$$

This transformation aligns the source domain data with the distribution of the target domain.

Next, we train a new regression model using the transported source domain data $\{ X_s^{'}, Y_s\}$, with the goal of improving predictive performance of the test data.

III. Data Acquisition & Data Analysis

A. Experiment Data Acquisition

To validate the model performance, we set up a data acquisition system to collect real stator current signals from a motor under different eccentricity conditions. An induction motor is modified to create controlled eccentricity: bearings were removed, and the rotor is supported by two custom-made mounting structures with a pair of new bearings on the extended rotor shaft. The motor's stator assembly is mounted on a linear stage, allowing horizontal position adjustments using two pairs of micrometers. Additionally, two pairs of displacement sensors are installed on the stator facing the air gap to measure the real air gap size in both horizontal and vertical directions while the motor is running [20]–[22]. A powder brake is connected to the test motor to serve as load.

In our experiment, data from three-phase current sensors were recorded at a 10 kHz sampling frequency for each SE and load condition. We created six SE levels in the horizontal direction while the motor was stationary: 7.1%, 16.5%, 31.1%, 42.5%, 47.5%, and 57.3%. These percentages represent the ratio of maximum air gap deviation to the nominal air gap size. The motor was tested under two load conditions: 0 N·m and 3.5 N·m.

B. Simulation Data Acquisition

Simulation data can be easily generated according to different fault and load conditions of the machine by the physics-based model described in Section II-A. For each load condition, we set the SE level from 5% to 70% for every 5%, run simulations, and obtain the stator current data with the same sampling frequency of 10 kHz. Therefore, much more simulation data can be obtained than the real data.

C. Data Analysis

For both simulation and real data, we segment 1,024 consecutive data points from the stator current in time-domain, corresponding to approximately 0.1 seconds of measurement at a 10 kHz sampling frequency, to investigate signal characteristics. Fig. 3 (a) and (c) display simulated data at an SE level of 40%, while Fig. 3 (b) and (d) show real data at an SE level of 42.5%, all recorded under no-load conditions. DQ transformation simplifies the current waveforms, enhancing the visibility of periodic characteristics and reducing complexity. Such an approach is especially valuable because real data often contains noise and irregularities, whereas simulated data tends to be cleaner and more periodic. Analyzing i_d signal allows for quantitative differentiation between simulated and real data patterns, capturing subtle phase and magnitude differences that may be less apparent through visual inspection alone.

IV. Numerical Test Setup

There different numerical experiments are designed to validate the effectiveness of Sim2Real DA, as shown in TABLE II. In practical applications, it is feasible to collect real data at normal condition and one faulty condition. We assume the data collected at SE levels 7.1% and 57.3% are available for training machine learning models. The rest of the measured data at SE levels 16.5%, 31.1%, 42.5%, and 47.5% are reserved as test data.

1) Real Data (Baseline): Only real measurement data at SE levels 7.1% and 57.3% are used for training.
2) Simulated Data without DA Algorithm: As a reference, we use simulated data generated by the physics-based

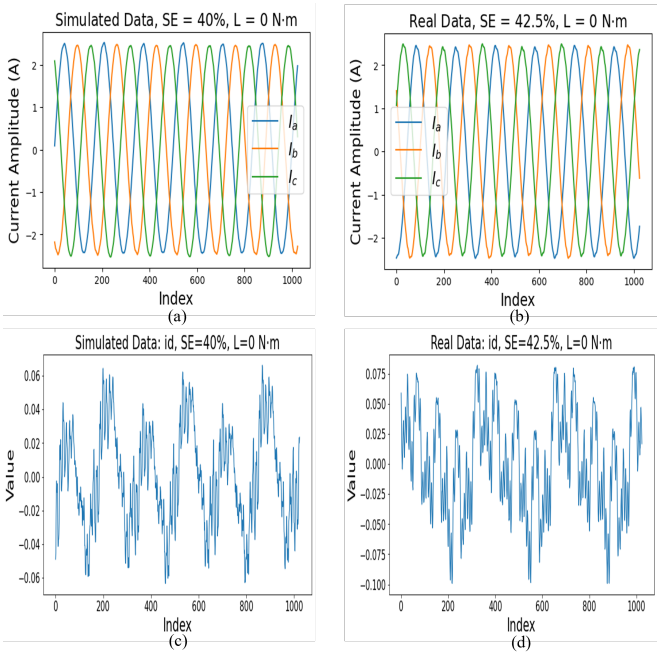

Fig. 2: Simulated data (a, c) at SE level of 40% and real data (b, d) at SE level of 42.5%, with load L=0 N·m, showing phase current signals in time-domain (top) and DQ-transformed i_d signal (bottom).

eccentricity model, with SE levels varying from 5% to 70%, as training data. No real data is used.

3) Simulated Data with DA Algorithm: Finally, we combine the available real data, and simulation data generated by the physics-based model (SE levels 5% to 70%), and apply Sim2Real DA process to align simulation data with real data and enhance the training data.

TABLE II: Experiment design

Data	Real Data (Baseline)	Simulated Data without Domain Adaptation	Simulated Data with Domain Adaptation
Training	Real Data (7.1, 57.3 levels)	All Simulated Data	All Simulated Data + Real Data (7.1, 57.3 levels)
Testing	Real Data (16.5, 31.1, 42.5, 47.5 levels)		

V. Results and Discussions

During training, a regression model is built using Support Vector Regression with a Radial Basis Function kernel (SVR-RBF) to predict SE levels based on extracted statistical features of i_d signal under two load conditions: no-load and 3.5 N·m load. During testing, model performance is evaluated using two key metrics: root-mean-squared-error (RMSE) and mean-absolute-error (MAE) on testing data. Both CORAL and OT methods are implemented separately for the DA process. The results, as shown in Table III, illustrate the prediction performance of load conditions: 0 and 3.5 N·m.

A. Simulated Data without Domain Adaptation

For no-load condition, both RMSE (6.76%) and MAE (5.76%) values are notably lower than those of the real data baseline (RMSE = 9.21%, MAE = 8.03%). This indicates that the simulated data captures features related to those in real data, even without DA. However, with 3.5 N·m load, the model prediction based on simulation data alone is much worse (with both RMSE and MAE over 20%) than the baseline result of using only real data, indicating large discrepancy between simulation and real data.

B. Simulated Data with Domain Adaptation (CORAL and OT)

Under the no-load condition, both CORAL and OT methods achieve RMSE and MAE values lower than the real data baseline. Specifically, the OT method reaches an RMSE of 5.65% (a 3.56% reduction) and an MAE of 5.28% (a 2.75% reduction) compared to the baseline (RMSE = 9.21%, MAE = 8.03%), showcasing their effectiveness in aligning simulated data with real data.

At the load condition of 3.5 N·m, lowest RMSE (5.59%) and MAE (4.77%) is obtained with OT method, outperforming the real data baseline and confirming OT's robustness in enhancing model accuracy. Additionally, the CORAL method also produces RMSE and MAE values lower than the baseline, showing that both DA methods improve performance under this load condition, though OT shows a more pronounced effect.

C. Data Visualization

Fig. 3 and Fig. 4 provide a visual distribution of SE level predictions under both no-load and on-load conditions. These violin plots allow us to observe the spread and concentration of prediction values across different experimental setups.

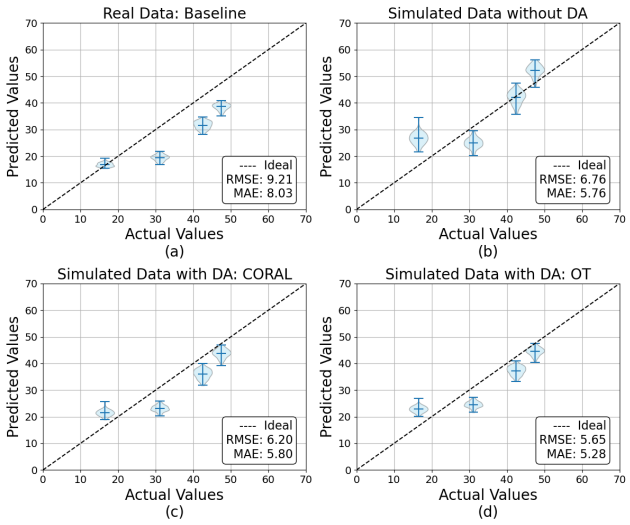

Fig. 3: Violin plot of SE level predictions on testing data under no-load condition using i_d signal to compare (a) real data: baseline, (b) simulated data without DA, (c) simulated data with DA:CORAL, and (d) simulated data with DA:OT.

TABLE III: Prediction Performance Comparison between Real Data and Simulated Data with and without Domain Adaptation

Load	Metric	Real Data (Baseline)	Simulated Data without Domain Adaptation	Simulated Data with Domain Adaptation:CORAL	Simulated Data with Domain Adaptation:OT
0	RMSE (%)	9.21	6.76	6.20	5.65
	MAE (%)	8.03	5.76	5.80	5.28
3.5 N·m	RMSE (%)	10.69	21.92	6.01	5.59
	MAE (%)	9.22	20.65	5.02	4.77

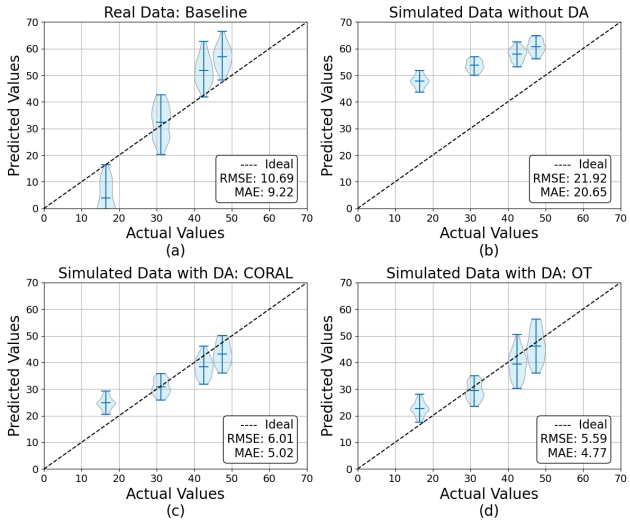

Fig. 4: Violin plot of SE level predictions on testing data at a consistent load condition of 3.5 N·m using i_d signal to compare (a) real data: baseline, (b) simulated data without DA, (c) simulated data with DA:CORAL, and (d) simulated data with DA:OT.

For no-load condition, simulated data without DA demonstrates a distribution relatively close to the actual values, confirming that some features of the simulated data naturally align with real data characteristics. With the application of DA, the distribution becomes even more refined.

For on-load condition, with real data only (Fig. 4(a)), the prediction has a large spread over true values. With simulation data alone (Fig. 4(b)) prediction result largely deviates from the true value, indicating the large mismatch between simulation and real data. The results with DA, for both CORAL and OT (Fig. 4(c) and (d)), show a narrower spread and a more centralized prediction distribution than the real data baseline and simulated data without DA. This concentrated prediction range indicates improved model stability with DA, particularly for the OT method, which exhibits the closest alignment with actual values.

For both load conditions, the OT method, in particular, shows a more concentrated distribution than CORAL, achieving a closer alignment to actual values and exhibiting a stable, centered prediction profile.

The effectiveness of the Sim2Real DA method can also be visualized with t-distributed stochastic neighbor embedding (t-SNE) plot in Fig. 5. During training, when comparing the differences between simulated data without DA and with DA, all simulated data is considered as source domain data, while all real data from the 7.1% and 57.3% SE levels is regarded as target domain data. Under both load: no-load and 3.5 N·m conditions, the distributions of source data and target data without DA are not well-aligned, as shown in Fig. 5 (a) and (c). However, with DA applied, the OT method aligns the target data more closely to the source data distribution, as illustrated in Fig. 5 (b) and (d).

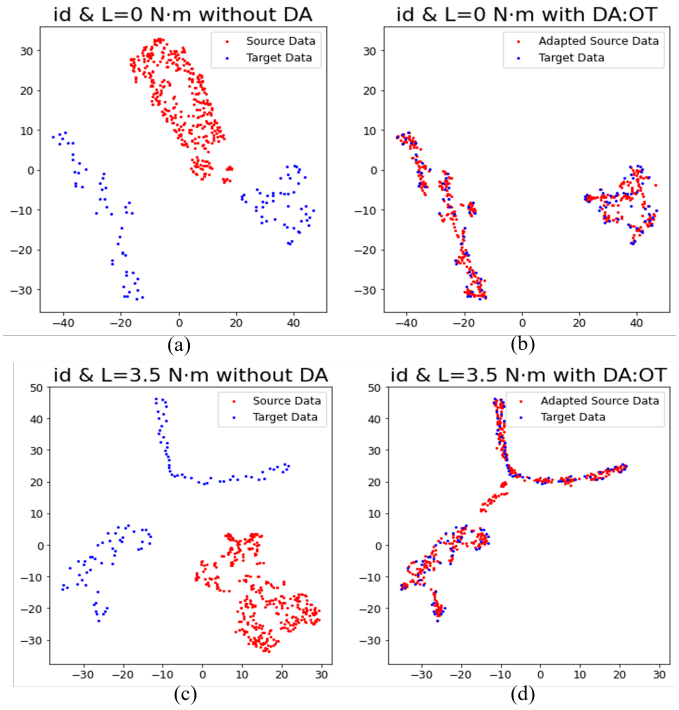

Fig. 5: t-SNE plot in 2d for i_d signal at: (a) no-load without DA, (b) no-load with DA, (c) 3.5 N·m without DA, and (d) 3.5 N·m with DA.

VI. Conclusion

This study presents a novel approach to motor eccentricity fault detection using simulation-to-reality domain adaptation. The proposed methodology, combining a physics-based eccentricity model, direct-quadrature transformation, and optimal transport for domain adaptation, demonstrates significant improvements in fault prediction accuracy compared to traditional methods relying solely on limited real data. The

results highlight the potential of leveraging simulated data to enhance fault detection capabilities, particularly in scenarios where acquiring extensive real fault data is impractical. The performance improvements observed under both no-load and loaded conditions validate the effectiveness of this simulation-to-reality approach in real-world applications. Future research can focus on extending this simulation-to-reality framework to address more complex motor fault scenarios and exploring advanced domain adaptation techniques to further bridge the gap between simulated and real data.

REFERENCES

[1] M. E. H. Benbouzid, "A review of induction motors signature analysis as a medium for faults detection," *IEEE transactions on industrial electronics*, vol. 47, no. 5, pp. 984–993, 2000.

[2] S. Nandi, H. A. Toliyat, and X. Li, "Condition monitoring and fault diagnosis of electrical motors—a review," *IEEE transactions on energy conversion*, vol. 20, no. 4, pp. 719–729, 2005.

[3] S. Zhang, S. Zhang, B. Wang, and T. G. Habetler, "Deep learning algorithms for bearing fault diagnostics—a comprehensive review," *IEEE Access*, vol. 8, pp. 29 857–29 881, 2020.

[4] J. Faiz and B. Ebrahimi, "Static eccentricity fault diagnosis in an accelerating no-load three-phase saturated squirrel-cage induction motor," *Progress in electromagnetics research B*, vol. 10, pp. 35–54, 2008.

[5] S. Nandi, S. Ahmed, and H. A. Toliyat, "Detection of rotor slot and other eccentricity related harmonics in a three phase induction motor with different rotor cages," *IEEE Transactions on Energy Conversion*, vol. 16, no. 3, pp. 253–260, 2001.

[6] K. N. Gyftakis and J. C. Kappatou, "A novel and effective method of static eccentricity diagnosis in three-phase psh induction motors," *IEEE Transactions on Energy Conversion*, vol. 28, no. 2, pp. 405–412, 2013.

[7] M. Akar, "Detection of a static eccentricity fault in a closed loop driven induction motor by using the angular domain order tracking analysis method," *Mechanical Systems and Signal Processing*, vol. 34, no. 1-2, pp. 173–182, 2013.

[8] H. A. Toliyat, M. S. Arefeen, and A. G. Parlos, "A method for dynamic simulation of air-gap eccentricity in induction machines," *IEEE transactions on industry applications*, vol. 32, no. 4, pp. 910–918, 1996.

[9] J. Faiz, B. M. Ebrahimi, B. Akin, and H. A. Toliyat, "Comprehensive eccentricity fault diagnosis in induction motors using finite element method," *IEEE Transactions on Magnetics*, vol. 45, no. 3, pp. 1764–1767, 2009.

[10] L. Zhou, B. Wang, C. Lin, H. Inoue, and M. Miyoshi, "Static eccentricity fault detection for psh-type induction motors considering high-order air gap permeance harmonics," in *2021 IEEE International Electric Machines & Drives Conference (IEMDC)*. IEEE, 2021, pp. 1–7.

[11] S. Zhang, F. Ye, B. Wang, and T. G. Habetler, "Semi-supervised bearing fault diagnosis and classification using variational autoencoder-based deep generative models," *IEEE Sensors Journal*, vol. 21, no. 5, pp. 6476–6486, 2020.

[12] ——, "Few-shot bearing fault diagnosis based on model-agnostic meta-learning," *IEEE Transactions on Industry Applications*, vol. 57, no. 5, pp. 4754–4764, 2021.

[13] A. Farahani, S. Voghoei, K. Rasheed, and H. R. Arabnia, "A brief review of domain adaptation," *Advances in data science and information engineering: proceedings from ICDATA 2020 and IKE 2020*, pp. 877–894, 2021.

[14] C. Liu and K. Gryllias, "Simulation-driven domain adaptation for rolling element bearing fault diagnosis," *IEEE Transactions on Industrial Informatics*, vol. 18, no. 9, pp. 5760–5770, 2021.

[15] Q. Wang, C. Taal, and O. Fink, "Integrating expert knowledge with domain adaptation for unsupervised fault diagnosis," *IEEE Transactions on Instrumentation and Measurement*, vol. 71, pp. 1–12, 2021.

[16] N. A. Al-Nuaim and H. Toliyat, "A novel method for modeling dynamic air-gap eccentricity in synchronous machines based on modified winding function theory," *IEEE Transactions on energy conversion*, vol. 13, no. 2, pp. 156–162, 1998.

[17] B. Wang, M. W. Albader, H. Inoue, and M. Kanemaru, "Induction motor eccentricity fault analysis and quantification with modified winding function based model," in *2022 25th International Conference on Electrical Machines and Systems (ICEMS)*. IEEE, 2022, pp. 1–6.

[18] B. Sun, J. Feng, and K. Saenko, "Return of frustratingly easy domain adaptation," in *Proceedings of the AAAI conference on artificial intelligence*, vol. 30, no. 1, 2016.

[19] N. Courty, R. Flamary, D. Tuia, and A. Rakotomamonjy, "Optimal transport for domain adaptation," *IEEE transactions on pattern analysis and machine intelligence*, vol. 39, no. 9, pp. 1853–1865, 2016.

[20] B. Wang, C. Lin, H. Inoue, and M. Kanemaru, "Topological data analysis for electric motor eccentricity fault detection," in *IECON 2022–48th Annual Conference of the IEEE Industrial Electronics Society*. IEEE, 2022, pp. 1–6.

[21] B. Wang, H. Inoue, and M. Kanemaru, "Motor eccentricity fault detection: Physics-based and data-driven approaches," in *2023 IEEE 14th International Symposium on Diagnostics for Electrical Machines, Power Electronics and Drives (SDEMPED)*. IEEE, 2023, pp. 42–48.

[22] B. Wang, C. Lin, H. Inoue, and M. Kanemaru, "Induction motor eccentricity fault detection and quantification using topological data analysis," *IEEE Access*, 2024.

BIBLIOGRAPHIES

Dai-Yan Ji received his B.S. degree in Electronic Engineering and M.S. degree in Communications Engineering from Feng Chia University, Taiwan, in 2009 and 2012, respectively. He is currently pursuing a Ph.D. degree in Mechanical Engineering with the University of Maryland, College Park, MD, USA. His research interests include machine learning, prognostics and health management, and industrial AI.

Bingnan Wang (M'12-SM'15) received his B.S. degree from Fudan University, Shanghai, China, in 2003, and Ph.D. degree from Iowa State University, Ames, IA, USA, in 2009, both in Physics. He has been with Mitsubishi Electric Research Laboratories (MERL), located in Cambridge, Massachusetts since then, and is now a Senior Principal Research Scientist. His current research focuses on electric machine analysis, design optimization, fault diagnosis with both physics modeling and machine learning techniques.

Hiroshi Inoue received B.E.and M.E. degree from The University of Tokyo, Japan, in 2012 and 2014, respectively. He has been a researcher at Mitsubishi Electric Corporation since April 2014. His primary research focus is motor malfunction diagnosis technology. He is currently a member of the Mitsubishi Electric Advanced Technology R&D Center, Electromechanical Systems Department.

Makoto Kanemaru received B.E., M.E. and Ph.D. degree from Tokyo Institute of Technology, Japan, in 2006, 2008 and 2011, respectively. He has been a researcher at Mitsubishi Electric Corporation since April 2011. From 2019 to 2020, he was a visiting scholar at Walter H. Shorenstein Asia-Pacific Research Center, Stanford University, USA. His primary research focus is motor malfunction diagnosis technology. He is currently a group manager in Electromechanical Systems Department of the Mitsubishi Electric Advanced Technology R&D Center.