공학석사학위논문

물류자동화 시스템의 고장진단을 위한 특징 분석 및 군집 적응형 네트워크 연구

Feature-analytic, Fleet-adaptive Network for Fault Diagnosis of Automated Material Handling Systems

2022 년 2 월

서울대학교 대학원

기계공학부

서 채 현

물류자동화 시스템의 고장진단을 위한 특징 분석 및 군집 적응형 네트워크 연구

Feature-analytic, Fleet-adaptive Network for Fault Diagnosis of Automated Material Handling Systems

지도교수 윤 병 동

이 논문을 공학석사 학위논문으로 제출함

2021 년 10 월

서울대학교 대학원

기계공학부

서 채 현

서채현의 공학석사 학위논문을 인준함

2021 년 12 월

위 원 장 : 김윤영 (인)

부위원장 : 윤병동 (인)

위 원 : 김도년 (인)

Abstract

Feature-analytic, Fleet-adaptive Network for Fault Diagnosis of Automated Material Handling Systems

Chaehyun Suh
Department of Mechanical Engineering
The Graduate School
Seoul National University

This paper proposes a Feature-analytic, Fleet-adaptive Network (FAFAN) for fault diagnosis of automated material handling systems (AMHSs) in semiconductor fabs. Constructing a fault-diagnosis model for a fleet of Overhead Hoist Transports (OHTs), which are the central part of AMHSs in semiconductor fabs, is challenging since the torque signals from different OHT units diverge from each other; further, the signals from many units consist of both labeled data and unlabeled data. To effectively deal with this situation, the proposed method learns fault-discriminative and OHT unit-domain-invariant features by selectively using pre-processed, multi-channel torque signals. Next, the approach independently extracts features from each channel and automatically learns the channel weights to leverage them, considering domain generalizability and the presence of fault signatures. The proposed method

consists of three main steps; 1) dividing the OHT dataset into a fully labeled source domain and a sparsely labeled target unit domain, 2) pre-processing front and rear torque signals into three-channel signals, and 3) extracting features to classify signals into normal, wheel fault, and gear fault states, while minimizing domain discrepancy through the use of semi-supervised domain adaptation. We demonstrate the effectiveness of the proposed method using data from 20 OHT units gathered from an actual industrial line, in numerous combinations of OHT unit domains, and different portions of target-domain-labeled data. The results of the validation verify that the proposed method is effective for fault diagnosis of a group of OHTs under insufficient label conditions and, further, that it provides physical evidence of the diagnosing conditions.

Keywords: Fault diagnosis
Overhead Hoist Transport (OHT)
Automated material handling system (AMHS)
Domain adaptation
Torque signal
Semi-supervised learning
Servo motor

Student Number: 2020-22707

Table of Contents

List of Tables

List of Figures

Nomenclatures

OHT	Overhead Hoist Transport
AMHS	Automated material handling systems
CNN	Convolutional neural network
SPMSM	Surface-mounted permanent magnet synchronous motor
PMSM	Permanent magnet synchronous motor
FAFAN	Feature-analytic, Fleet-adaptive Network
MAF	Moving average filte
MMD	Maximum mean discrepancy
BN	Batchnormalization
MP	Maxpooling
GAP	Global average pooling
FC	Fully-connected layer
SF	Softmax
SIG	Sigmoid
$S-only$	Source only
STL	Source Target Labeled
$STL-CSA$	Source Target Labeled CSA loss
$STL-CSA-MMD$	Source Target Labeled CSA loss with Maximum Mean Discrepancy

Chapter 1. Introduction

1.1 Research motivation

Overhead Hoist Transport (OHT) systems are an essential part of automated material-handling systems in semiconductor factories. These transport systems play a key role in the transportation of wafers, by autonomously moving on guide rails installed on the factory ceiling [1]. Since the driving parts of OHTs are subjected to continuous high loading, and because OHT systems usually run continuously, their components, such as their wheels and gearbox, are likely to degrade rapidly. If any fault occurs in an OHT unit, other units running on the same railway are at risk of being affected by the immobile unit, thereby putting the whole production line at risk. Therefore, monitoring the conditions and diagnosing the fault status of every OHT unit is crucial in maintaining a high wafer fabrication yield in a factory. To help maintain reliable operation of automated systems and their subcomponents, fault diagnosis methods have been developed, using signals measured from components, such as vibration, current, and torque signals [2][3][25][26]. In particular, fault diagnosis methods using current or torque signals obtained from electric motor controllers are scalable to many machines in an actual line, since the acquisition of these signals requires no additional sensors. For OHTs, two-channel torque signals from the controllers of dual-motor systems that are gathered for command purposes are also available for condition monitoring. However, it is difficult to generalize a fault diagnosis model for use with many OHT units, since there are discrepancies in the distribution of signals obtained from different units. Also, even though enough labeled data can be obtained from some units for fault monitoring, labeled data from many OHT units may be insufficient, and the portion

of unlabeled data can be large due to practical issues that arise in labeling. Thus, to construct a model that is applicable to not only a single OHT unit, but also to a group of OHTs in an actual production line, deviations in the two-channel torque signals gathered from the units must be overcome using unlabeled data and the small number of labeled data.

To address the challenges of constructing a common fault diagnosis model using unlabeled and sparsely labeled data from different machines, transfer learning, domain adaptation, and semi-supervised learning methods have been suggested [5][6][7][8][9][27]. Transfer learning has shown decent performance for transferring knowledge from the informative source domain; it is usually executed using sparsely labeled, target-domain data for fine-tuning the trained model. Domain-adaption methods, which are a category of transfer learning, learn the feature spaces to overcome the distributional discrepancy between domains in unsupervised or semi-supervised settings, for use in the target domain. Common shallow domain-adaptation methods aim to learn the mapping between fixed-source and target-domain features. In industrial scenarios, these methods mostly rely on hand-engineered input representations. Recently, deep domain-adaptation methods that use deep architectures, such as a convolutional neural network (CNN), have shown superior performance, since they can learn domain-invariant feature spaces autonomously by backpropagation. The main branches of these methods are based on adversarial training and a distance metric. Guo et al. [16] developed a domain-adaptation model for fault diagnosis of bearings using unlabeled vibration signals from three different testbeds, and Li et al. [8]developed a generalized model for adapting real fault data to artificial bearing fault data utilizing the pseudo-label

technique.

However, since each channel of a multi-channel torque signal obtained from the dual motors in an OHT can contain different physical characteristics and meanings, the conventional CNN architecture utilized in the prior work described in the literature has a limitation in extracting uniqueness in multi-channel input settings, since the features from each channel are combined from the input layer with the same kernel size [17][18]. Moreover, it is hard to quantify the relative contributions of each channel for conducting fault diagnosis. In addition, leveraging this information in the training procedure is difficult. These difficulties can prevent effective model training and proper interpretation of the black-box model.

1.2 Research scope

This paper proposes a new method, a Feature-analytic, Fleet-adaptive Network, that performs semi-supervised OHT unit-domain adaptation, considering different characteristics and differing fault signatures in each channel of the torque signals. The proposed method extracts features from a channel-independent CNN architecture and automatically learns to weigh them to perform domain adaptation and classification. Moreover, by analyzing the trained weight scores for each channel, investigations of the contributions of each channel of the torque signal to the given task can be performed. First, in the proposed method, the source and target OHT unit domains are divided according to the portion of labeled data in each unit. Second, the two-channel front and rear motor torque signals are pre-processed into three-channel torque signals (*Residual*, *Sum*, and *Difference*) and standardized within each

channel. Third, features for each input channel are extracted through the use of channel-independent CNN and are combined with weight to minimize classification loss and semantic alignment loss (using labeled data) and to maximize domain discrimination loss (using unlabeled data). In this paper, we limit the scope of the fault modes of the OHT to gear and wheel faults.

1.3 Dissertation Layout

This paper is organized as follows. In section 2, the basics of OHTs and the characteristics of dual-motor system torque signals are introduced. Next, the proposed method is described in detail in section 3. Section 4 describes validation of the proposed method using data from a real-world OHT line. Finally, the conclusion and the contribution of the paper are summarized, and the future work is suggested, in section 5.

Chapter 2. Background

In this section, first, the configuration of an OHT is explained in section 2.1. Then, the characteristics of the control torque signals obtained from the dual-motor system in an OHT are investigated in section 2.2.

2.1 Overhead Hoist Transport (OHT)

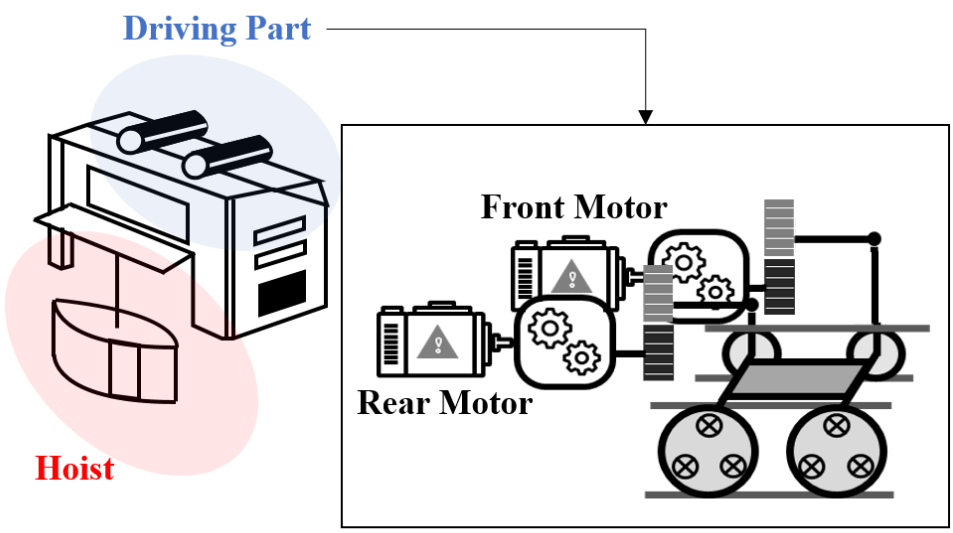

Figure 1. Overall configuration of Overhead Hoist Transport (OHT)

An OHT consists of two parts: a hoist part that carries wafers and a driving part that maneuvers the whole system on a railway attached to the ceiling. The driving part is steered by two speed-controlled surface-mounted permanent magnet synchronous motors (SPMSMs) that are aligned along a guide rail. OHTs run at a constant speed on a regular basis along the straight sections of the closed-curve railway system. Thus, to reduce variability and randomness that arise from the drive commands for acceleration and deceleration, torque signals from the front and rear SPMSMs are extracted from the constant-speed regions for all OHT units in a system.

In this paper, we aim to build a model that classifies the health condition of an OHT as normal, wheel fault, or gear fault; these states were chosen since wheel and gear faults represent a large proportion of OHT fault modes. The wheel fault mode occurs due to wheel wear that is caused by the continuous high friction between the urethane wheel and the railway. The gear fault mode consists of backlash and oil leak in the gearbox, which are caused by continuous load transmissions.

2.2 Characteristics of the control torque signals of OHTs

Fault diagnosis can be performed using the torque signal of a motor, since the fault-related periodic components at the load are transmitted from the load torque to the electromagnetic torque T_E produced in the air gap of the motor. For a permanent magnet synchronous motor (PMSM), T_E can be expressed as [19]:

$$T_E = J\frac{dw}{dt} + \sum_{n=1}^{\infty} T_n \cos(2\pi f_n t + \varphi_n) + T_{L,C} + T_0 \quad (1)$$

where J is the moment of inertia of the system driven by the motor, ω is the mechanical angular velocity of the rotor, T_E is electromagnetic torque, $T_{L,C}$ is the constant part of the load torque, T_0 is no-load torque, and T_n, f_n, φ_n are the amplitude, frequency, and phase angle of the harmonic component of the periodic fault in the load torque, respectively. In a constant-speed control scheme, (1) can be summarized as:

$$T_E = T_{E,C} + \sum_{n=1}^{\infty} T_n \cos(2\pi f_n t + \varphi_n) \tag{2}$$

where $T_{E,C}$ is the constant part of the electromagnetic torque. $T_{E,C}$ is affected mainly by the command of the controller to output a certain amount of DC torque to obtain the desired speed. Assuming that the change that results from the addition of the payload is negligible, based on the weight of the total OHT system, this part could be affected by the non-periodic mechanical faulty component at the load torque [20]. In the case where a dual-motor system is the OHT driving part, the faulty information from any part of either the front or rear motor's load propagates through the drivetrain and consequently affects the mutual characteristics of $T_{E,C}$ for both the front and rear motors [21][22]. Therefore, informative pre-processing of the torque signals to highlight certain parts related to the fault can enhance the fault diagnosis performance.

Since dozens of OHTs run on a single railway, it is important to monitor every unit, since any single faulty unit can affect the whole process. However, constructing a

torque-signal-based fault diagnosis feature or a model that is applicable to many OHTs is challenging, as the signals obtained from each unit diverge from each other due to configuration parameter changes and the inherent randomness of the control signal. Furthermore, labeled data – especially for fault modes – are not sufficient for many OHT units in real production lines, since abnormal events rarely happen and the labeling process is labor-intensive. Therefore, these challenges have to be addressed in order to conduct fault diagnosis of OHTs in an actual operation line.

Chapter 3. Proposed method

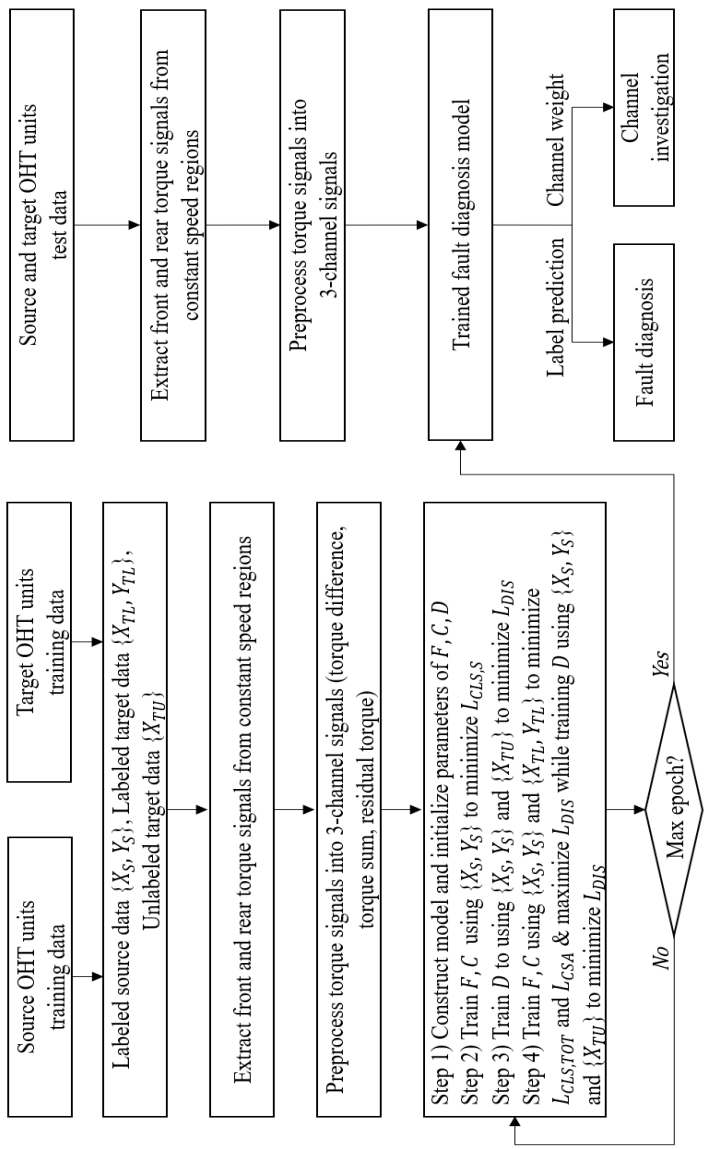

Figure 2. Overall flowchart of the proposed method

This section describes the details of the proposed Feature-analytic, Fleet-adaptive Network (FAFAN) for fault diagnosis of OHTs using semi-supervised unit-domain adaptation. Fig. 2 shows the overall flowchart of the proposed FAFAN method. The architecture of the proposed method is introduced in section 3.1, followed by an explanation of the training steps for the model, which are described in section 3.2.

3.1 Configuration of the proposed FAFAN method

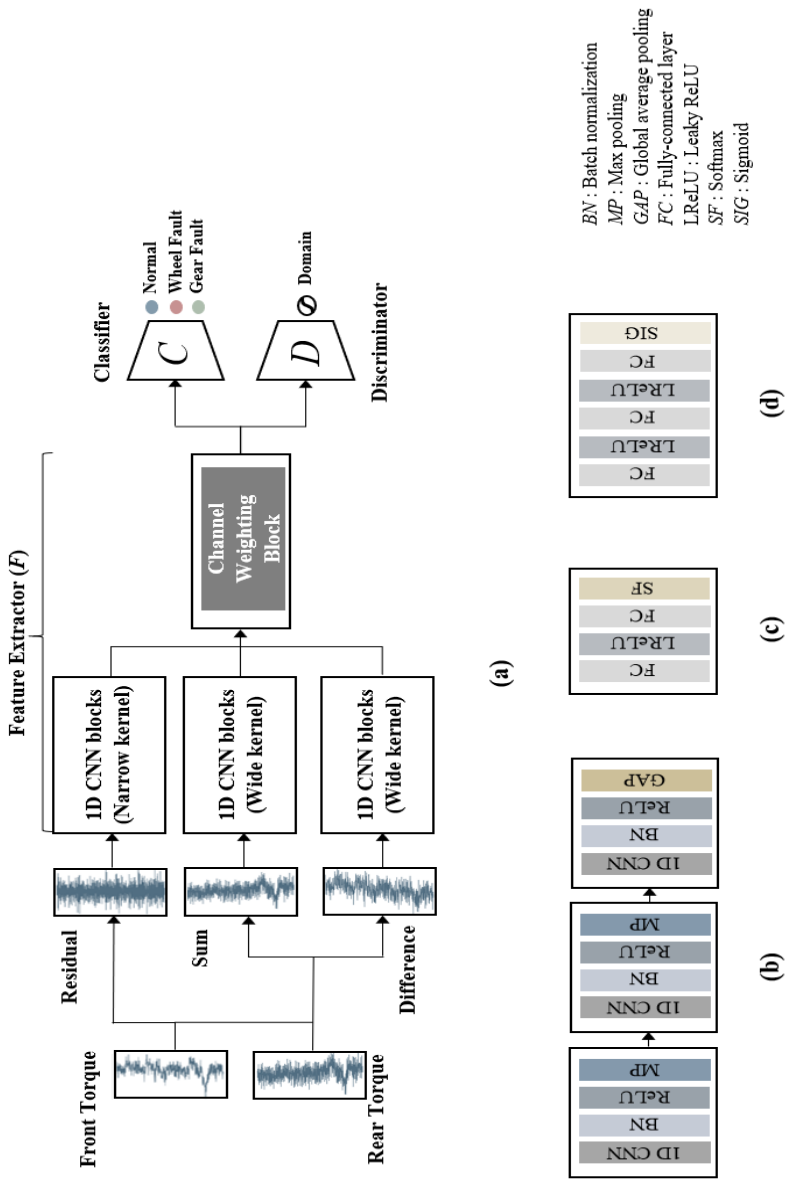

Figure 3. Architecture of FAFAN: (a) overall FAFAN process, (b) 1D CNN block, (c) Condition classifier, (d) Domain discriminator

The proposed method is designed to deal with multi-channel torque signals as it pursues the objective of learning OHT unit-domain invariant latent features, while achieving high fault classification accuracy using both labeled and unlabeled data. The proposed architecture develops conventional CNN's autonomous feature extraction ability to independently extract and leverage different fault characteristics of each input torque channel. The overall architecture of the proposed FAFAN approach is shown in Fig. 3; it consists of four parts, which are described in detail in the following subsections.

3.1.1 Pre-processing module

First, constant-speed regions are identified the using motor encoder signals and the front and rear torque signals of an OHT are extracted from these regions. Then, these two-channel torque signals are pre-processed into three-channel torque signals, including the *Residual*, *Sum*, and *Difference*, to highlight the different parts of the torque signal that are affected by a fault. To obtain the *Residual*, a moving average filter (MAF) that averages three consecutive time steps is applied to the raw front torque signal and the filtered signal is subtracted from the original signal. This results in a torque signal where the control-related part ($T_{E,C}$ in (1)) is minimized and contains load torque oscillations in the high-frequency regions. Furthermore, to consider the relation between the two SPMSMs in faulty load conditions, we define *Sum* as the summation of the front and rear torque signals and *Difference* as the subtraction of the rear torque from the front torque. These linear combinations of

front and rear torques contain control-related parts and some low-frequency load torque oscillations. Since the control command is affected by the fault as it achieves constant speed, the total power output of the OHT driving part is characterized by the sum, and the power output asymmetry in the dual-motor system is characterized by the *Difference*, accordingly [22]. After pre-processing, each of the three-channel inputs is normalized independently to form a distribution of zero mean and a unit standard deviation, considering all training data.

3.1.2 Feature extractor *F*: Channel-independent CNN

A CNN is constructed in three paths independently to extract unique features from each channel. The convolutional layers in the CNN are assigned to each of the input channels and are not combined with one another until the global average pooling layer. Since high-frequency components dominate the *Residual* and low-frequency components dominate the *Sum* and *Difference*, the kernel size that represents the receptive field size for acquiring information from the raw time-series input is set as relatively small for the *Residual* and relatively large for the *Sum* and *Difference* [23]. This whole process can induce the deep neural network to examine different regions in the raw input and automatically extract meaningful features for the given task.

3.1.3 Feature extractor *F*: Channel-weighting block

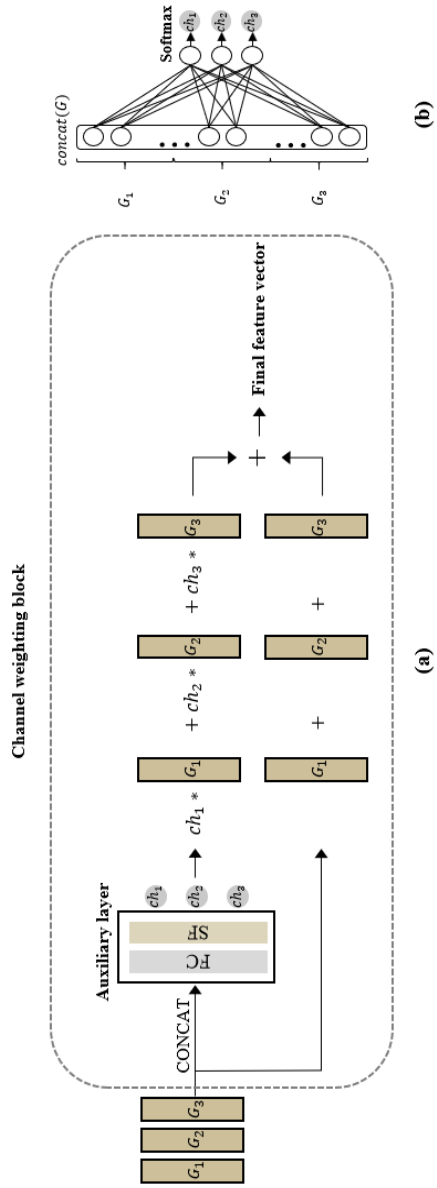

Figure 4. Configuration of the channel-weighting block: (a) Channel-weighting block, (b) Auxiliary layer

To quantify the relative contributions of each channel for the given tasks of classification and domain adaption and to reflect this information in the training procedure, a channel-weighting block that contains an auxiliary layer is adopted to automatically learn the weights of each channel, as shown in Fig. 4. These weights (ch_i (3)) are calculated from the auxiliary layer; they take a concatenation of the global average pooled feature vectors (G_i in (3)) from each channel as the input and output of the softmax function values. Every node in a fully connected layer combines all values of the previous layer to consider the relationships among features from each input channel and finally from among each of the three channels [14]. The outputs are three positive values that sum up to 1; each of them is multiplied by each channel's global-average-pooled feature vector. The linear summation of each channel's global-average-pooled feature vector is passed on the other side, which forms a residual connection to maintain the original information from each channel and to facilitate the training procedure. The final feature vector is ($F(X)$ in (3)), which is obtained by summing these two vectors. The result is used to form a latent feature space for the feature extractor part. The equations of this process are as follows:

$$ch_i = \frac{e^{\sum_k w_{i,k} * concat(G)_{i,k} + b_i}}{\sum_{j=1}^{3} e^{\sum_k w_{j,k} * concat(G)_{j,k} + b_j}}, i = 1,2,3 \;\&\; \sum_{i=1}^{3} ch_i = 1$$
$$F(X) = (ch_1 * G_1 + ch_2 * G_2 + ch_3 * G_3) + (G_1 + G_2 + G_3) = \sum_{i=1}^{3} (ch_i + 1) * G_i \quad (3)$$

where G_i is the global-average-pooled vector of the i^{th} input channel, $w_{i,k}$ is the k^{th} weight, b_i is the bias of i^{th} input channel, and X is the input data.

3.1.4 Task module: Condition classifier *C* & Domain discriminator *D*

After the features from each channel are linearly summed using their weight, the resultant feature vector is fed into two fully connected layers to perform domain discrimination and condition classification tasks. In particular, the feature extractor (part 2 & 3 of the architecture) is used to learn to deceive the domain discriminator in an adversarial way, to learn the OHT unit-domain-invariant features.

3.2 Model training procedures

3.2.1 Train *F* and *C* to classify the condition

F and *C* are trained to minimize the source classification loss $L_{CLS,S}$ using only source data $\{X_S,\ Y_S\}$ that consist of fully labeled normal, wheel fault, and gear fault data. Since the target domain is sparsely labeled, the latent feature space is first constructed using an informative source domain, which will be used as a decent initial base for later transformation. The classification loss used in this step is a cross-entropy loss for 3-class classification. $L_{CLS,S}$ is expressed as follows:

$$L_{CLS,S}(\theta_F,\theta_C)=\frac{1}{n_S}\sum_{n=1}^{n_S}CE(C(F(X_{S,i})),Y_{S,i})=-\frac{1}{n_S}\sum_{n=1}^{n_S}\sum_{j=1}^{3}1[Y_{S,i}=j]\log(\frac{e^{C(F(X_{S,i}))_j}}{\sum_{k=1}^{3}e^{C(F(X_{S,i}))_k}}) \quad (4)$$

where θ_F and θ_C are the parameters of F and C respectively, n_S is the number of source domain samples, and CE is cross-entropy loss. The term $[Y_{S,i}=\mathrm{j}]$ returns 1 when the equality holds.

3.2.2 Train *D* using to discriminate the OHT unit domain

D is trained to discriminate the source and target domain OHT units using the learned F. The binary domain label is assigned as 1 for the source domain and 0 for the target domain to distinguish the two domains. The source data X_S, the unlabeled target data X_{TU}, and the corresponding domain label Y_D, are used to minimize the domain discrimination loss L_{DIS}, given the fixed feature space from the previous step. This step is crucial since developing a well-performing D can improve the result of the adversarial learning in the next step by intensively challenging F to modify the feature space to fool this strong opponent. L_{DIS} is expressed as follows:

$$\begin{aligned}L_{DIS}(\theta_D)&=\frac{1}{n_S+n_{TU}}\sum_{n=1}^{n_S+n_{TU}}BCE(D(F(X_i)),Y_{D,i})\\&=-\frac{1}{n_S+n_{TU}}\sum_{n=1}^{n_S+n_{TU}}(Y_{D,i}\log(D(F(X_i)))+(1-Y_{D,i})\log(1-D(F(X_i))))\end{aligned} \quad (5)$$

where θ_D denotes the parameters of D, n_{TU} is the number of unlabeled target domain samples, and BCE is the binary cross-entropy loss.

3.2.3 Train *F*, *C*, and *D* to learn generalized feature representation for the source and target domains

The F and C are trained to minimize the total classification loss $L_{CLS,TOT}$ and contrastive semantic alignment loss L_{CSA} using source data $\{X_S,\ Y_S\}$ and labeled target data $\{X_{TL},\ Y_{TL}\}$. $L_{CLS,TOT}$ consists of $L_{CLS,S}$ and target classification loss $L_{CLS,T}$, which is expressed as follows:

$$L_{CLS,T}(\theta_F,\theta_C)=\frac{1}{n_{TL}}\sum_{i=1}^{n_{TL}}CE(C(F(X_{TL,i})),Y_{TL,i})=-\frac{1}{n_{TL}}\sum_{n=1}^{n_{TL}}\sum_{j=1}^{3}1[Y_{TL,i}=j]\log(\frac{e^{C(F(X_{TL,i}))_j}}{\sum_{k=1}^{3}e^{C(F(X_{TL,i}))_k}})$$
$$L_{CLS,TOT}=\alpha_S L_{CLS,S}+\alpha_T L_{CLS,T} \quad (6)$$

where n_{TL} is the number of labeled target domain samples; α_S and α_T are the weights to balance the source and target classification loss functions. The L2-norm contrastive semantic alignment loss L_{CSA} measures the distance between the source and target domain latent features of the same label (first term in the right hand side of (7)) and of the different label (first term in the right hand side of (7)); this loss is minimized to place the features of the same label near each other and to regulate the features of the different label to be far from each other at the margin (it is set to be 1 in this paper). L_{CSA} is expressed as follows:

$$L_{CSA}(\theta_F,\theta_C)=\frac{1}{n_{pairs}}\Big(\sum_{i,samelabel}\|F(X_{S,i})-F(X_{TL,i})\|^2+\sum_{i,difflabel}\max(0,margin-\|F(X_{S,i})-F(X_{TL,i})\|)^2\Big) \quad (7)$$

where n_{pairs} is the number of source and labeled target domain data pairs to derive the distance. At the same time, the latent features are trained to make D fail to identify from which domain the OHT unit comes, by maximizing the domain discrimination loss L_{DIS} using X_S and X_{TU}, as in the previous. D also learns to update the loss function to make a balance in the adversarial learning procedure. This whole step utilizes given sparsely labeled target data to make them fit in the pre-trained feature space in the first step, which ultimately results in a generalized diagnosis model applicable to many OHT units. The total loss function $L_{CLS,TOT}$ and the optimal parameters are expressed as follows:

$$\begin{aligned}
&L_{TOT}(\theta_F,\theta_C,\theta_D)=\alpha_{CLS}L_{CLS,TOT}(\theta_F,\theta_C)+\alpha_{CSA}L_{CSA}(\theta_F,\theta_C)-\lambda L_{DIS}(\theta_F,\theta_D)\\
&(\theta_F,\theta_C)=\arg\min_{\theta_F,\theta_C}L_{TOT}(\theta_F,\theta_C,\theta_D)\\
&(\theta_D)=\arg\max_{\theta_D}L_{TOT}(\theta_F,\theta_C,\theta_D)
\end{aligned} \quad (8)$$

where $\widehat{\theta_F}$, $\widehat{\theta_C}$ and $\widehat{\theta_D}$ are the optimal parameters of F, C, and D respectively. α_{CLS}, α_{CSA}, and λ are the weights to balance the classification, contrastive semantic alignment, and domain adversarial learning processes.

Chapter 4. Experimental validation

In this section, the performance of the proposed method is evaluated using data obtained from OHTs running in an actual line. In section 4.1, the dataset used for experimental studies is described. Then, the comparison methods are outlined in section 4.2, followed by experimental results, which are provided in section 4.3.

4.1 Dataset description

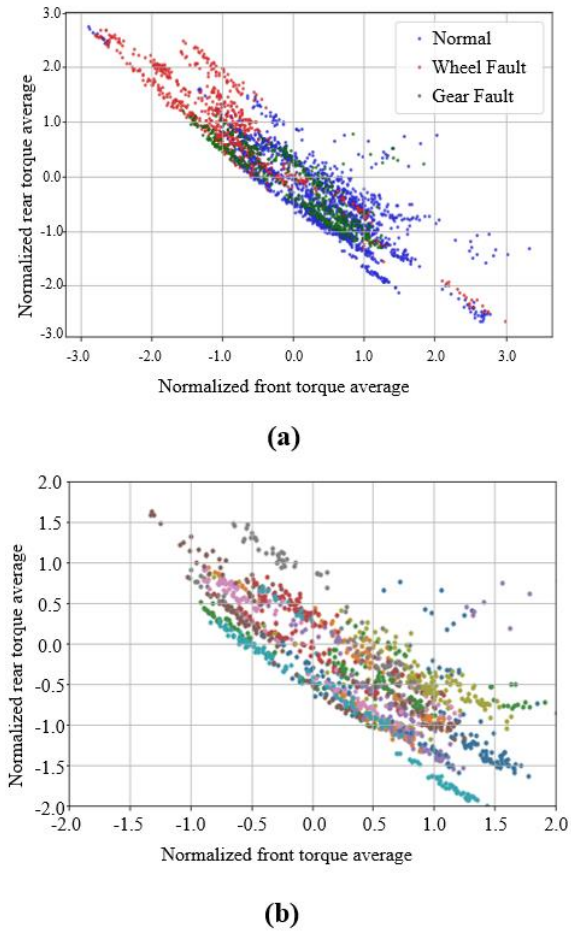

Figure 5. Scatter plot of average values of the front and rear torque signals: (a) all classes, (b) only normal class

A dataset consisting of torque signals from 20 OHT units from actual semiconductor factories is used for validation of the proposed method. The torque

signals are sampled with a frequency of 1 kHz from the front and rear motors in the driving part of an OHT. Three health conditions of normal, wheel fault, and gear fault are included in the dataset.

The averages of the normalized front and rear torque data of some OHT units are shown in Fig. 5. It is hard to determine decision boundaries to classify the health conditions in Fig. 5 (a) since the intra-class variance is large for all classes and the inter-class variance is small, especially for the normal and gear fault states. The large intra-class variance is due to the distributional shift between OHT units, as visualized by the different colors in Fig. 5 (b); the small inter-class variance implies the challenge in extracting distinguishable features using torque signals. To build a generalized fault diagnosis model, these difficulties have to be addressed using a ratio of a small amount of labeled data to a large amount of unlabeled data, for most OHT units.

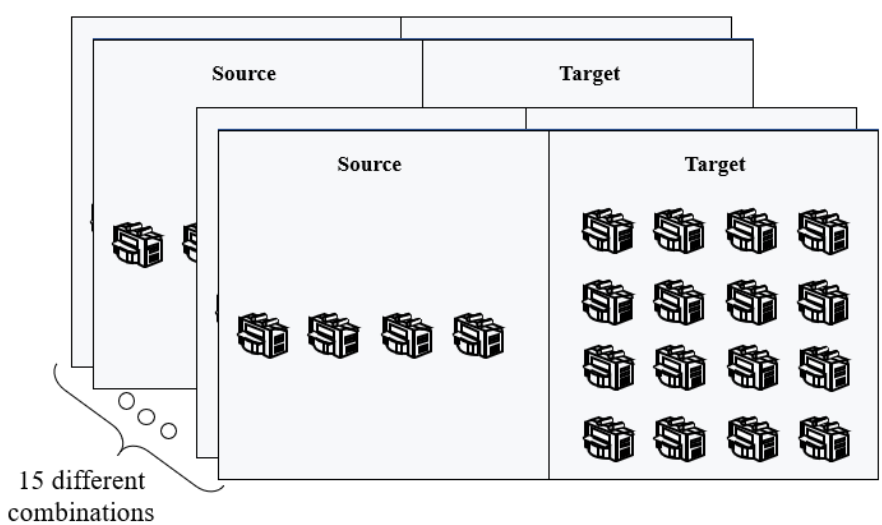

Figure 6. Source and target domain OHT configuration

To verify the effectiveness of the proposed method for unit-wise domain adaptation, given data are divided into the source and target domains. The source domain consists of four fully labeled OHT units; the target domain consists of 16 sparsely labeled OHT units. A total of 15 different combinations of these domains are considered by randomly splitting the OHT units.

4.2 Description of the comparison methods

4.2.1 Source only (S-only)

For the S-only method, the architecture of the Feature-analytic, Fleet-adaptive Network is trained to minimize $L_{CLS,S}$ using only the source-domain data $\{X_S,\ Y_S\}$. The fault diagnosis performance of this method is directly related to the distance between the source and target domains; the lower the performance, the larger the domain gap.

4.2.2 Source Target Labeled (STL)

For the STL method, labeled target-domain data $\{X_{TL},\ Y_{TL}\}$ are added to S-only to minimize $L_{CLS,TOT}$. This method shows the necessity of contrastive semantic alignment, which aims to minimize L_{CSA}.

4.2.3 Source Target Labeled CSA loss (STL-CSA)

For the STL-CSA method, the same data that were used in STL are used again to train the STL-CSA model; however, the optimization objective is to minimize $\alpha_{CLS}L_{CLS,TOT} + \alpha_{CSA}L_{CSA}$. L_{CSA} is added to reduce the domain gap and increase the inter-class variance. This method is compared to show the effect of domain adaptation using unlabeled data.

4.2.4 Source Target Labeled CSA loss with Maximum Mean Discrepancy (STL-CSA-MMD)

Maximum mean discrepancy (MMD) is widely used to quantify the discrepancy between the source and the target domains, using statistical distance. The MMD loss also utilizes unlabeled data, added to form $\alpha_{CLS}L_{CLS,TOT} + \alpha_{CSA}L_{CSA} + \alpha_{MMD}L_{MMD}$.

$$
\begin{aligned}
&L_{MMD} = MMD^2(F(X_S), F(X_{TU})) \\
&= \frac{1}{n_S^{\,2}}\sum_{i,j=1}^{n_S} k(F(X_{S,i}), F(X_{S,j})) - \frac{1}{n_S n_{TU}}\sum_{i=1}^{n_S}\sum_{j=1}^{n_{TU}} k(F(X_{S,i}), F(X_{TU,j})) + \frac{1}{n_{TU}^{\,2}}\sum_{i,j=1}^{n_{TU}} k(F(X_{TU,i}), F(X_{TU,j})) \\
&k(x, x') = \exp\left(-\frac{\| x - x' \|^2}{2\sigma^2}\right), \sigma^2 = 0.01, 0.1, 1, 10
\end{aligned}
\tag{9}
$$

MMD is the empirical maximum mean discrepancy estimated using the kernel trick. The kernel is chosen as the Gaussian radial basis function (RBF)[24] and, since the choice of the bandwidth parameter σ^2 greatly affects the performance, it is set to four different values, ranging from 0.1 to 10, to enhance robustness.

4.3 Experimental settings

All experiments were conducted on a Windows 10 computer with an Intel i7-9700 CPU (3.00 GHz) and an NVIDIA GeForce RTX 2080 SUPER GPU (3072 CUDA cores, 8GB GDDR6 memory). The code was written in Python 3.7 and the PyTorch deep learning framework was used for model training and testing. The experiments for all methods were conducted with the learning rate of 0.0009, batch size of 16, and an exponential learning decay rate of 0.99. λ is set as $\frac{2}{1+\exp(-10p)} - 1$, where $\mathrm{p} = \frac{\text{current epoch}}{\max \text{epoch}}$ .The hyper-parameters of the deep architecture (feature extractor, condition classifier, and the domain discriminator) are listed in Table 1; the comparison methods share the same feature extractor and condition classifier as those used for the proposed FAFAN.

Table 1. Network configuration and parameters of the proposed method

Layer		Parameters	Output [Channel, Length]
Input		Window size = 1000	[3, 1000]
Feature extractor (*F*) - CNN	CONV 1	Kernel size = 8 (*Residual*) / 200 (*Sum*, *Difference*), Channels = 18, Stride = 1, Activation = ReLU, Batch normalization	[18, 993] (*Residual*) / [18, 801] (*Sum*, *Difference*)
	MP 1	Kernel size = 2, Stride = 2	[18, 496] (*Residual*) / [18, 400] (*Sum*, *Difference*)
	CONV 2	Kernel size = 8, Channels = 36, Stride = 1, Activation = ReLU, Batch normalization	[36, 489] (*Residual*) / [36, 393] (*Sum*, *Difference*)
	MP 2	Kernel size = 2, Stride = 2	[36, 244] (*Residual*) / [36, 196] (*Sum*, *Difference*)
	CONV 3	Kernel size = 8, Channels = 72, Stride = 1, Activation = ReLU, Batch normalization	[72, 237] (*Residual*) / [72, 189] (*Sum*, *Difference*)
	GAP	-	[1, 72]
Feature extractor (*F*) - Channel-weighting block	CONCAT	Concatenate three [1, 72] GAPs from *Residual, Sum*, and *Difference*	[1, 216]
	FC	Node = (216, 3), Output = Softmax	[1, 3]
Condition classifier (*C*)	FC 1	Node = (72, 36), Activation = LeakyReLU (0.2)	[1, 36]
	FC 2	Node = (36, 3), Output = Softmax	[1, 3]
Domain discriminator (*D*)	FC 1	Node = (72, 64), Activation = LeakyReLU (0.2)	[1, 64]
	FC 2	Node = (64, 32), Activation = LeakyReLU (0.2)	[1, 32]
	FC 3	Node = (32, 1), Output = Sigmoid	[1, 1]

For the performance analysis, the training and test data are split with the ratio of 75:25, in a way that any pair of training and test data do not come from the same constant-speed region. In particular, for fair validation, the training and test data for a target OHT unit are divided in an inductive setting, where the test data with label information are not used in the training procedure in any form.

4.4 Results and discussion

4.4.1 Performance analysis

The fault diagnosis accuracies for the target domain test set in each target labeled data condition are calculated using S-only, STL, STL-CSA, STL-CSA-MMD, and the proposed method, as shown in Fig. 7. The accuracies are calculated for 15 different source and target OHT unit combinations; the mean accuracy is represented as the height of the bar, and the black vertical line centered at the top of the bar represents the 95% confidence interval of the corresponding accuracy.

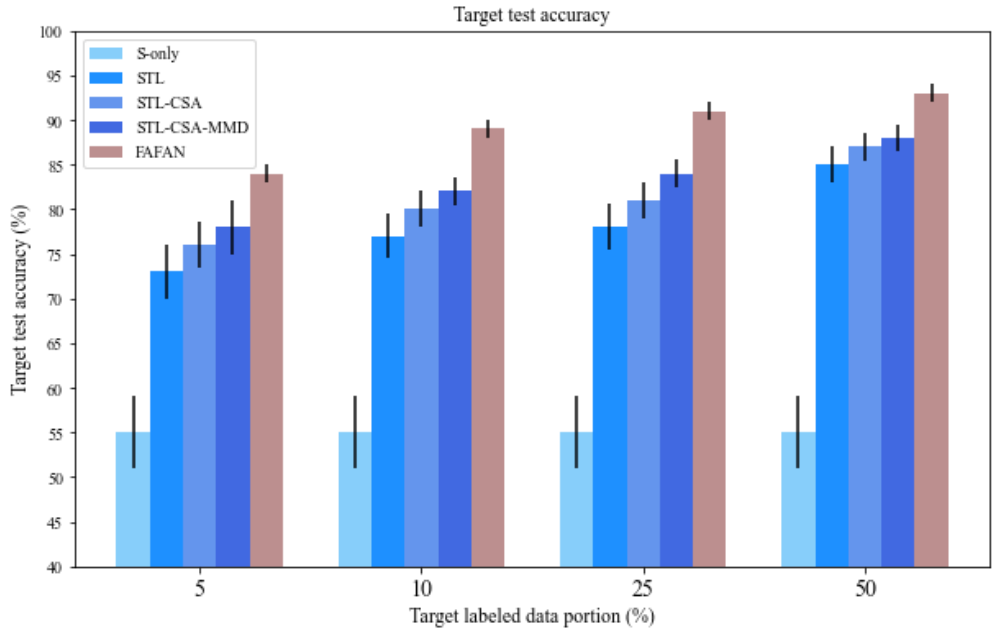

Figure 7. Comparisons of fault diagnosis accuracies for target domain test data

Table 2. Fault diagnosis accuracies (%) for target domain test data

Method	**Target labeled percent (%)**			
	5	**10**	**25**	**50**
S-only	55±4	55±4	55±4	55±4
STL	73±3	77±3	78±3	85±2
STL-CSA	76±3	80±2	81±2	87±2
STL-CSA-MMD	78±3	82±2	84±2	88±2
FAFAN	**84±1**	**89±1**	**91±1**	**93±1**

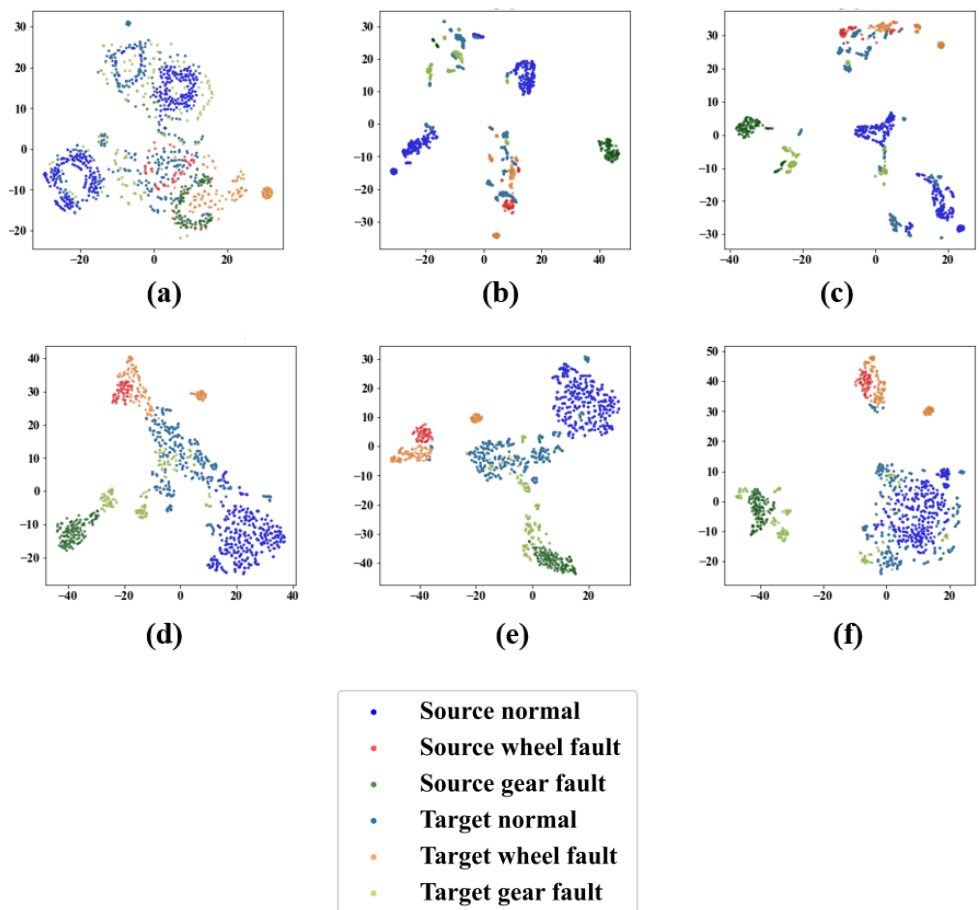

Figure 8. t-SNE visualization of the latent feature space: (a) Raw time series, (b) S-only, (c) STL, (d) STL-CSA, (e) STL-CSA-MMD, (f) FAFAN

Fig 8 shows 2D visualization of the high-dimensional deep latent space of the methods in the 5% target label case, using t-Distributed Stochastic Neighbor Embedding. As seen in Fig. 8 (a), the input data is complex, and the decision boundaries for classification are impossible to determine. The feature space of S-only and STL, as shown in Fig. 8 (b) and (c), indicate that the gap between the domains is large and even the features within the domain are divergent. Thus, by adding L_{CSA} as for STL-CSA, the features are more organized; however, the decision boundaries are imperfect, since the unlabeled data in the target domain are not considered in the training procedure, as shown in Fig. 8 (d). By adding MMD loss, the discrepancy of the source and target domain was slightly reduced; however, the effect in fault diagnosis is limited. The FAFAN approach utilizes the domain discriminator in an adversarial way, such that it eventually fails to identify to which domain the feature belongs. In Fig. 8 (f) the features from different domains but with the same label are clustered near each other. The features in Fig. 8 (d) can be distinguished easily into the source and target domains by a simple domain decision boundary; however, the features in Fig. 8 (e) successfully interrupt the domain discriminator in determining the domain decision boundary.

4.4.2 Input channel investigation

The softmax outputs of the channel-weighting block are analyzed to investigate whether the physical behavior is reflected in the learning of the proposed model. Each of the softmax outputs ch_1, ch_2, and ch_3 corresponds to the learned weight of each input channel *Residual*, *Sum*, and *Difference.*

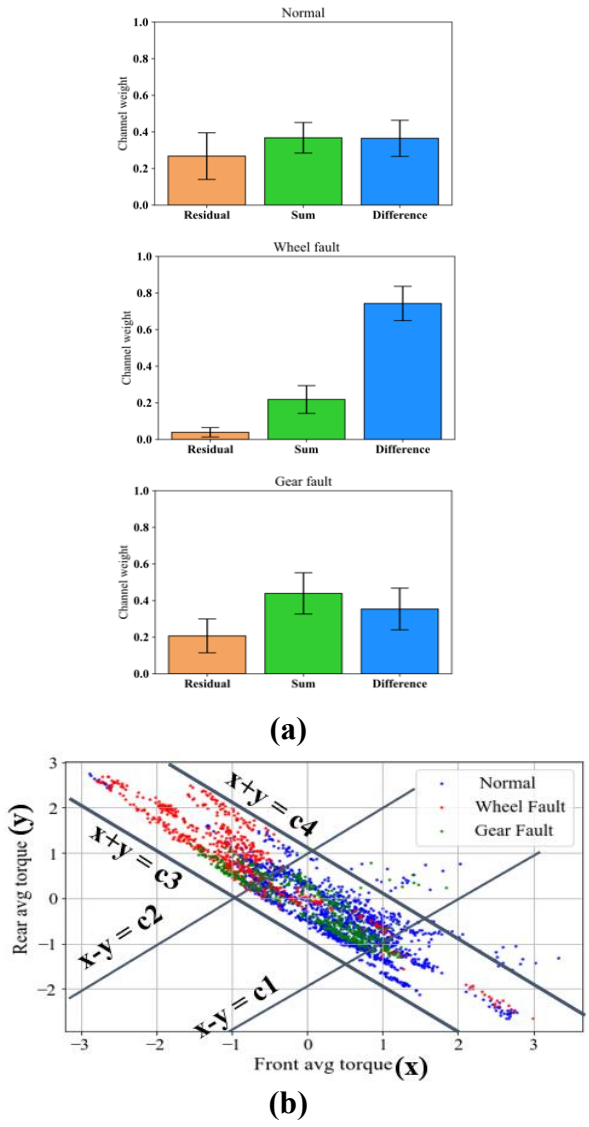

Figure 9. (a) Statistics of weights in the channel weighting block, by health condition, (b) Torque sum and difference line representation in the scatter plot of average values

When FAFAN predicts each health condition given new test data, the features from each input channel are weighted by the trained auxiliary layer in the channel-weighting block. The mean and standard deviations of these weights are calculated in the target domain and are shown in Fig. 9 (a). When predicting the normal class, each channel is evenly weighted, as seen in Fig. 9 (a). Estimation of the wheel fault mode requires the model to highlight the *Difference.* Also, in the torque average scatter plot in Fig. 9 (b), many wheel fault data points can be distinguished using the *Difference*. When a wheel fault occurs, the unbalance of the output energy between the front and rear motor tends to increase, as compared to the normal condition. This physical phenomenon is reflected in the training procedures of the proposed method, as it extracts more class-distinguishable and domain-invariant features than the mere average of the front and rear torque. For the gear-fault mode, the channel-weighting block highlights slightly more on the *Sum* than it does for the other input channels. The proposed method identifies the discriminability of the gear fault compared to the normal condition, using *Sum*, and utilizes this information to extract deep features.

Chapter 5. Conclusion

5.1 Summary

This paper proposes a Feature-analytic, Fleet-adaptive Network (FAFAN) that can quantify the relative importance of each multi-channel torque signal for conducting semi-supervised OHT unit-domain adaptation. FAFAN consists of pre-processing module, feature extractor that is a cascade of channel-independent CNN and channel-weighting block, and task module of condition classifier and domain discriminator. FAFAN is trained in three steps to learn fault-discriminative but OHT unit-domain-invariant latent feature space. FAFAN's fault diagnosis capability was validated using real line OHTs torque data and it was compared with other methods such as S-only that uses model trained only with source domain data, STL that uses model trained with both source and target labeled data, STL-CSA that considers contrastive semantic alignment loss additional to STL, and STL-CSA-MMD that considers MMD additionally to STL-CSA to minimize discrepancy between source and target domains using unlabeled target domain data. Consequently, the proposed FAFAN was superior to other existing methods using datasets of various portions of target domain labeled data and source and target OHT unit domain combinations. The analysis of channel-weighting block after training provided evidence of diagnosing faulty conditions in the physical context of the dual-motor torque signals.

5.2 Contribution

The proposed method offers the following primary contributions:

Contribution 1: Unified fault diagnosis model for a fleet of AMHS

The proposed method considers the signal discrepancy among different machines of the same type to construct a unified fault diagnosis model using a fully labeled source domain and a sparsely labeled target domain.

Contribution 2: Fault diagnosis method validated using data from actual manufacturing lines

The proposed method is validated using torque data obtained from dozens of OHTs running in actual semiconductor manufacturing lines, in contrast to most fault diagnosis research, which is conducted only with laboratory-scale data.

Contribution 3: Enhancing explainability of the deep model using channel weights

The proposed method's architecture learns the relative weight score of each input time-series channel to enhance feature learning for fault diagnosis and domain adaptation and to additionally investigate which channel was relatively considered in the given task. This eventually enhances explainability of the deep model in estimating health states in the physical context such as total power output and power output asymmetry.

Contribution 4: Control signal based fault diagnosis method for general application

The proposed method utilizes the command control torque signal from a surface-mounted permanent magnet synchronous motor (SPMSM), which does not require any additional sensors, to diagnosis mechanical faults such as wheel and gear fault. Thus, the proposed framework is widely applicable to diverse industrial scenarios where control signals are already available.

5.3 Future work

Future work will examine methods to improve the pre-processing and network configuration to enhance the diagnosis performance and physical interpretability of the deep neural network. To improve the pre-processing of two-channel torque signals, filters that reduce the noise of the signal can be implemented and advanced methods that capture the subtle relation between two-channel torque can be figured out. To improve the deep neural network configuration, parameter optimization using grid search or Bayesian optimization can be used. Further, extended applications of the proposed method to different industrial systems will be conducted in future studies such as industrial robots and other motor-driven mechanical systems.

References

[1] A. Zhakov et al., “Application of ANN for Fault Detection in Overhead Transport Systems for Semiconductor Fab,” IEEE Trans. Semicond. Manuf., vol. 33, no. 3, pp. 337–345, Aug. 2020, doi: 10.1109/TSM.2020.2984326.

[2] C. H. Park et al., “Drive-tolerant Current Residual Variance (DTCRV) for Fault Detection of a Permanent Magnet Synchronous Motor Under Operational Speed and Load Torque Conditions,” IEEE Access, pp. 1–1, 2021, doi: 10.1109/ACCESS.2021.3068425.

[3] C. H. Park, H. Kim, J. Lee, G. Ahn, M. Youn, and B. D. Youn, “A Feature Inherited Hierarchical Convolutional Neural Network (FI-HCNN) for Motor Fault Severity Estimation Using Stator Current Signals,” Int. J. Precis. Eng. Manuf. Technol., Oct. 2020, doi: 10.1007/s40684-020-00279-3.

[4] Y. Kim, J. Park, K. Na, H. Yuan, B. D. Youn, and C. soon Kang, “Phase-based time domain averaging (PTDA) for fault detection of a gearbox in an industrial robot using vibration signals,” Mech. Syst. Signal Process., vol. 138, Apr. 2020, doi: 10.1016/j.ymssp.2019.106544.

[5] S. Kim et al., “A Semi-Supervised Autoencoder With an Auxiliary Task (SAAT) for Power Transformer Fault Diagnosis Using Dissolved Gas Analysis,” IEEE Access, vol. 8, pp. 178295–178310, Sep. 2020, doi: 10.1109/access.2020.3027830.

[6] K. Saito, D. Kim, S. Sclaroff, T. Darrell, and K. Saenko, “Semi-supervised

domain adaptation via minimax entropy," in Proceedings of the IEEE International Conference on Computer Vision, Oct. 2019, vol. 2019-October, pp. 8049–8057, doi: 10.1109/ICCV.2019.00814.

[7] X. Li, W. Zhang, and Q. Ding, "Cross-domain fault diagnosis of rolling element bearings using deep generative neural networks," IEEE Trans. Ind. Electron., vol. 66, no. 7, pp. 5525–5534, Jul. 2019, doi: 10.1109/TIE.2018.2868023.

[8] Y. Li, Y. Song, L. Jia, S. Gao, Q. Li, and M. Qiu, "Intelligent Fault Diagnosis by Fusing Domain Adversarial Training and Maximum Mean Discrepancy via Ensemble Learning," IEEE Trans. Ind. Informatics, vol. 17, no. 4, pp. 2833–2841, Apr. 2021, doi: 10.1109/TII.2020.3008010.

[9] Q. Wang, G. Michau, and O. Fink, "Missing-Class-Robust Domain Adaptation by Unilateral Alignment," IEEE Trans. Ind. Electron., vol. 68, no. 1, pp. 663–671, Jan. 2021, doi: 10.1109/TIE.2019.2962438.

[10] Y. Lecun, Y. Bengio, and G. Hinton, "Deep learning," Nature, vol. 521, no. 7553. Nature Publishing Group, pp. 436–444, May 27, 2015, doi: 10.1038/nature14539.

[11] J. Donahue et al., "DeCAF: A Deep Convolutional Activation Feature for Generic Visual Recognition," arXiv preprint arXiv:1310.1531, Oct. 2013.

[12] X. Wang, L. Li, W. Ye, M. Long, and J. Wang, "Transferable Attention for Domain Adaptation." in Proceedings of the AAAI Conference on Artificial

Intelligence, vol. 33, no. 1, pp. 5345–5352, Jul. 2019, doi: 10.1609/aaai.v33i01.33015345.

[13] C. Zhang, Q. Zhao, and Y. Wang, “Transferable attention networks for adversarial domain adaptation,” Inf. Sci. (Ny)., vol. 539, pp. 422–433, 2020, doi: 10.1016/j.ins.2020.06.016.

[14] S. Woo, J. Park, J.-Y. Lee, and I. S. Kweon, “CBAM: Convolutional Block Attention Module,” arXiv preprint arXiv:1807.06521, Jul. 2018.

[15] Y. Ganin and V. Lempitsky, “Unsupervised domain adaptation by back-propagation,” in Proc. 32nd Int. Conf. Mach. Learn., vol. 37, pp. 1180–1189, Jul. 2015.

[16] L. Guo, Y. Lei, S. Xing, T. Yan, and N. Li, “Deep Convolutional Transfer Learning Network: A New Method for Intelligent Fault Diagnosis of Machines with Unlabeled Data,” IEEE Trans. Ind. Electron., vol. 66, no. 9, pp. 7316–7325, Sep. 2019, doi: 10.1109/TIE.2018.2877090.

[17] Y. Zheng, Q. Liu, E. Chen, Y. Ge, and J. L. Zhao, “Time series classification using multi-channels deep convolutional neural networks,” in Proc. 15th Int. Conf. Web-Age Inf. Manag., Macau, China, pp. 298–310, Jun. 2014.

[18] S. Shao, R. Yan, Y. Lu, P. Wang, and R. X. Gao, “DCNN-Based multi-signal induction motor fault diagnosis,” IEEE Trans. Instrum. Meas., vol. 69, no. 6, pp. 2658–2669, Jun. 2020, doi: 10.1109/TIM.2019.2925247.

[19] Y. Yao, Y. Li, and Q. Yin, “A novel method based on self-sensing motor drive

system for misalignment detection," Mech. Syst. Signal Process., vol. 116, pp. 217–229, Feb. 2019, doi: 10.1016/j.ymssp.2018.06.030.

[20] C. Verucchi, J. Bossio, G. Bossio, and G. Acosta, "Misalignment detection in induction motors with flexible coupling by means of estimated torque analysis and MCSA," Mech. Syst. Signal Process., vol. 80, pp. 570–581, Dec. 2016, doi: 10.1016/j.ymssp.2016.04.035.

[21] K. Cho and D.-H. Lee, "Design and Position Control of Overhang-Type Rail Mover Using Dual BLAC Motor," Energies, vol. 14, no. 4, p. 1000, Feb. 2021, doi: 10.3390/en14041000.

[22] W. Chen, Y. Wu, R. Du, Q. Chen, and X. Wu, "Speed tracking and synchronization of a dual-motor system via second order sliding mode control," Math. Probl. Eng., vol. 2013, 2013, doi: 10.1155/2013/919837.

[23] B. Zhu et al., "Learning Environmental Sounds with Multi-scale Convolutional Neural Network," Proc. Int. Jt. Conf. Neural Networks, vol. 2018-July, 2018, doi: 10.1109/IJCNN.2018.8489641.

[24] A. Gretton, K. M. Borgwardt, M. J. Rasch, B. Schölkopf, and A. Smola, "A kernel two-sample test," J. Mach. Learn. Res., vol. 13, pp. 723–773, 2012.

[25] J. U. Ko, J. H. Jung, M. Kim, H. B. Kong, J. Lee, and B. D. Youn, "Multi-task learning of classification and denoising (MLCD) for noise-robust rotor system diagnosis," Comput. Ind., vol. 125, p. 103385, 2021, doi: 10.1016/j.compind.2020.103385.

[26] J. H. Jung, M. Kim, J. U. Ko, H. B. Kong, B. D. Youn, and K. H. Sun, “Label-based, Mini-batch Combinations Study for Convolutional Neural Network Based Fluid-film Bearing Rotor System Diagnosis,” Comput. Ind., vol. 133, p. 103546, 2021, doi: 10.1016/j.compind.2021.103546.

[27] Y. Deng, D. Huang, S. Du, G. Li, C. Zhao, and J. Lv, “A double-layer attention based adversarial network for partial transfer learning in machinery fault diagnosis,” Comput. Ind., vol. 127, p. 103399, 2021, doi: 10.1016/j.compind.2021.103399.

국문 초록

물류자동화 시스템의 고장진단을 위한 특징 분석 및 군집 적응형 네트워크 연구

서울대학교 공과대학
기계공학부 대학원
서 채 현

본 논문은 반도체 공장의 물류자동화 시스템 (AMHS)의 고장 진단을 위한 특징 분석 및 군집 적응형 네트워크 (FAFAN)를 제안한다. 반도체 공장 AMHS의 핵심인 천장 반송 시스템 (OHT) 군집에 대한 고장 진단 모델을 구축하는 것은, 각 OHT 호기별로 토크 신호의 편차가 존재하기 때문에 어렵다. 또한, 많은 호기에서 취득되는 신호는 정상/고장 레이블이 있는 데이터와 레이블이 없는 데이터로 구성되어 있다. 이러한 상황에서 제안된 방법은, 전처리된 다채널 토크 신호를 활용하여 고장을 진단함과 동시에 OHT 호기 도메인에 대한 일반적인 특징을 학습한다. 특히, 전처리된 입력 채널에서 특징을 독립적으로 추출하고 도메인 일반화 가능성과 고장 진단의 정보량을 모델 학습 과정에 활용하기 위해 채널 가중치를 자동으로 학습한다. 제안된 방법은 1) 레이블이 있는 데이터로만 구성된 소스 도메인과, 레이블이 있는

데이터가 매우 적은 타겟 도메인으로 OHT 데이터세트를 나누는 단계, 2) 전면 및 후면 토크 신호를 3채널 신호로 전처리하는 단계, 그리고 3) 준지도 도메인 적응을 활용하여 OHT 호기 도메인 간의 신호 편차를 최소화함과 동시에 정상, 바퀴 결함 및 기어 결함 상태로 분류하는 특징을 추출하는 단계로 구성된다. 실제 산업 현장에서 수집된 20개의 OHT 호기 데이터에 대해, 많은 OHT 호기 도메인 조합 및 타겟 도메인 레이블 데이터의 다양한 비율 조합을 활용하여 제안된 방법의 효과를 입증한다. 검증 결과, 제안된 방법이 불충분한 레이블 데이터 조건에서 OHT 군집의 고장 진단에 효과적이며, 나아가 진단 결과의 물리적 근거를 제공함을 확인하였다.

주요어: 고장 진단
천장 반송 시스템
물류자동화 시스템
도메인 적응
토크 신호
준지도 학습

학 번: 2020-22707