DOI: 10.61782/fa.2025.0022

FORUM ACUSTICUM EURONOISE 2025

# A NEW FEATURE EXTRACTION FRAMEWORK FOR SOURCE-FREE DOMAIN ADAPTATION IN BEARING FAULT DIAGNOSIS

**Sakthisandron Arunagiri** **Sitesh Kumar Mishra** **Piyush Shakya***
Engineering Asset Management Group, Department of Mechanical Engineering,
Indian Institute of Technology Madras, Chennai-36, India

## ABSTRACT

Condition monitoring of machine elements is critical in industrial applications, drawing significant attention to cross-domain bearing fault diagnosis. However, data privacy concerns and storage limitations pose significant challenges. To address these issues, source-free domain adaptation networks, consisting of a feature extractor and a fault classifier, have been developed. In these frameworks, the parameters of both components are trained using only source-domain data. During testing on target-domain data, the extractor's parameters are initialized with the learned source-domain parameters and allowed to adapt, while the classifier's parameters remain fixed. Although the proposed approach effectively tackles privacy and storage concerns, it falls short in providing a unique feature representation for each domain, limiting its generalization capability. The present work seeks to overcome the mentioned limitation by employing auto-regressive models to generate unique feature representations for individual domains. Additionally, self-organizing maps are utilised to explore similarities across domains and various fault types. The robustness of the proposed approach is validated through extensive experiments, demonstrating its effectiveness in improving cross-domain fault diagnosis performance.

**Keywords:** *bearing diagnosis, source free domain adaptation, auto-regressive model, self-organizing map.*

**Corresponding author*: *pshakya@iitm.ac.in.*


## 1. INTRODUCTION

In modern industrial systems, predictive maintenance plays a crucial role in ensuring safety and minimizing downtime. Significant research efforts have been dedicated to developing data-driven methods for early fault detection [1, 2], which can assist in effective decision-making. However, real-world implementation faces several challenges, including large data storage requirements, data privacy concerns, and the increased difficulty of domain adaptation when applying models across different operating conditions. Recent studies [3,4] have attempted to address these issues by developing source-free domain adaptation frameworks. This work aims to tackle these challenges by introducing a novel feature extraction approach that provides a unique representation of the dataset, ensuring more distinctive and domain-specific feature learning.

A source-free domain adaptation approach for bearing fault diagnosis begins with feature extraction from the source-domain bearing data followed by training a classifier using the corresponding fault labels. Notably, training is performed exclusively on the source data, without access to target labels. Once trained, the same feature extraction procedure is applied to the target-domain bearing data, and the classifier is expected to accurately classify faults in the target domain despite distributional differences.

Mathematically, the labelled source-bearing dataset is represented as $D_S = \{x_i^S, y_i^S\}_{i=1}^{N_S}$, where $x_i^S$ and $y_i^S$ denote the vibration signal and the corresponding fault label for the $i^{th}$ sample, respectively. The unlabelled target-bearing dataset is given by $D_T = \{x_i^T\}_{i=1}^{N_T}$, where $N_S$ and $N_T$ represent the number of samples in the source and target datasets, respectively. Since the source and target data distributions may differ, the key challenge is to learn a mapping function: $f : x_i \rightarrow y_i$ using only the source data, such that it can effectively classify faults in the target domain despite the absence of target labels.

## 2. PROPOSED METHOD

### 2.1 Overview

The proposed method consists of three main stages: feature extraction, feature representation, and classification. In the feature extraction stage, the vibration signal is modelled as the output of an Auto-Regressive (AR) process [5]. The AR model parameters and the residual variance are extracted as features to ensure a unique representation for each sample. In the feature representation stage, the extracted features are structured using Self-Organizing Maps (SOMs) [6]. The transformation organizes the samples based on similarity, improving feature discrimination and clustering. In the classification stage, a deep neural network (DNN) is trained on the SOM-structured feature representations of the source domain. Once trained, the parameters of both the SOM and the classifier are frozen. The same feature extraction and representation process is then applied to the target domain data. The structured representations of the source and target samples are clustered independently. To facilitate unsupervised domain adaptation, the target clusters are mapped to the closest corresponding source clusters. Finally, the mapped target samples are classified using the trained DNN. An overview of the proposed method is illustrated in Fig. 1.

### 2.2 Mathematical Framework of the Proposed Method

#### 2.2.1 *Feature Extraction Using AR Model*

All vibration signals in the dataset are modelled as the output of an AR process of order five as

$$x(k) = \sum_{j=1}^{5} \alpha_j x(k-j) + \epsilon(k) \tag{1}$$

where $\epsilon(k)$ is white noise with variance $\sigma^2$, and $\alpha_j$ represents the $j^{th}$ coefficient. These coefficients, along with the residual noise, constitute the feature vector for each vibration signal, expressed as $\mathbf{f} = [\alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_5, \sigma^2]$. Consequently, the transformed dataset is given by $\bar{D}_S = \{\mathbf{f}_i^S, y_i^S\}_{i=1}^{N_S}$ and $\bar{D}_T = \{\mathbf{f}_i^T\}_{i=1}^{N_T}$.

#### 2.2.2 *Structured Representation via SOM*

The extracted features are mapped to a structured 2D latent space using a SOM consisting of a $20 \times 20$ neuron lattice, with $\mathbf{w_{ij}}$ representing its weights. The iterative process for updating these weights is as follows:

- Find the Best Matching Unit (BMU):

$$(i^*, j^*) = \arg\min_{(i,j)} \|\mathbf{f}_i - \mathbf{w}_{ij}\|$$

- Update neuron weights:

$$\mathbf{w}_{ij} \leftarrow \mathbf{w}_{ij} + \eta . h(i,j)(\mathbf{f}_i - \mathbf{w}_{ij})$$

  where $\eta$ is the learning rate, and $h(i,j) = e^{\frac{(i-i^*)^2+(j-j^*)^2}{2\beta^2}}$ is the Gaussian neighborhood function with $\beta$ as the neighborhood radius.
- Repeat until the weights converge.
  After convergence, the structured representations from SOM$_S$ and SOM$_T$ are denoted as $\mathbf{w}_{ij}^S$ and $\mathbf{w}_{ij}^T$, respectively.

#### 2.2.3 *Training DNN-Based Classifier*

The updated weights $\mathbf{w}_{ij}^S$ from SOM$_S$ are stacked to form $\mathbf{z}^S$, which serves as input to the DNN for training its weights and biases. The DNN architecture is mathematically expressed as:

$$\hat{y}_S = \sigma_3(\mathbf{W}_3(\sigma_2(\mathbf{W}_2(\sigma_1(\mathbf{W}_1\mathbf{z}^S + \mathbf{b}_1)) + \mathbf{b}_2) + \mathbf{b}_3)) \tag{2}$$

where $\mathbf{W}_*, \mathbf{b}_*$, and $\sigma_*$ denote the weights, biases, and ReLU activation functions, respectively. Training is performed for 2000 epochs to minimize the cross-entropy loss between $\hat{y}_S$ and the true source labels using the Adam optimizer.

#### 2.2.4 *Unsupervised Domain Adaptation and Classification*

After training, the weights $\mathbf{w}_{ij}^S$ of SOM$_S$ and the parameters of the DNN are frozen while adapting to the target dataset. The procedure for aligning $\mathbf{w}_{ij}^T$ to $\mathbf{w}_{ij}^S$ is as follows:

- Apply K-means clustering to the neurons of SOM$_S$ and SOM$_T$:

$$C^S = \{c_{ij}^S\}_{(i,j)\in \text{SOM}_S}, \quad C^T = \{c_{ij}^T\}_{(i,j)\in \text{SOM}_T}$$

  where $c_{ij}^S$ and $c_{ij}^T$ represent the cluster assignments for each neuron.

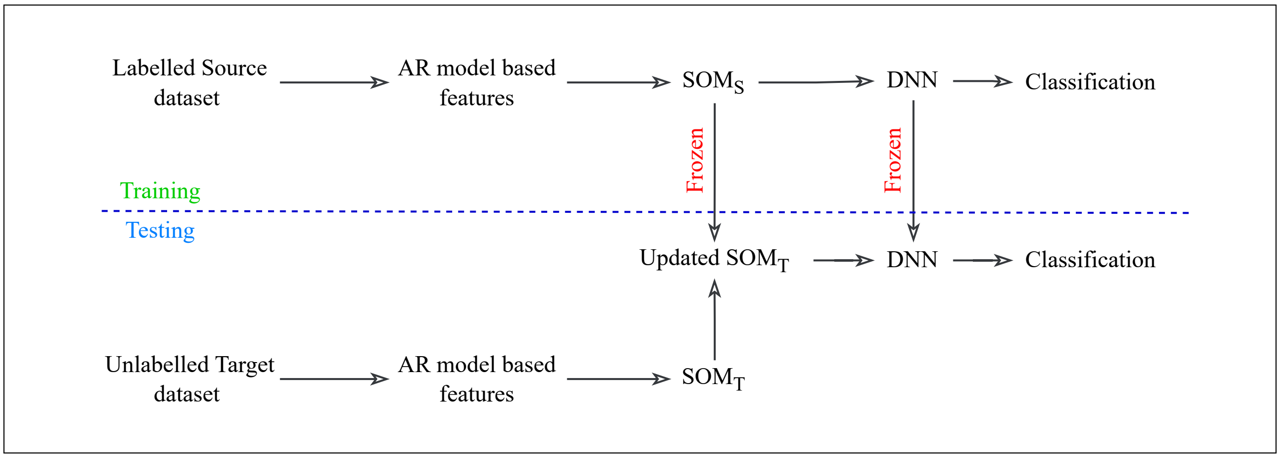


**Figure 1**. Overview of the proposed method.

- Compute the centroid of source clusters:

$$\mu_k^S = \frac{1}{|C_k^S|} \sum_{(i,j)\in C_k^S} \mathbf{w}_{ij}^S$$

  where $C_k^S$ is the set of neurons assigned to cluster $k$.
- Update target weights iteratively:

$$\mathbf{w}_{ij}^T \leftarrow \lambda \mathbf{w}_{ij}^T + (1-\lambda)\mu_k^S$$

Finally, the updated $\mathbf{w}_{ij}^T$ is fed into the DNN for fault classification:

$$\hat{y}_T = DNN(\mathbf{w}_{ij}^T) \tag{3}$$

## 3. DISCUSSION

### 3.1 Dataset and Tasks

The proposed method is validated on both same-domain and cross-domain tasks using bearing vibration data from the CWRU dataset [7] and the IMS dataset [8]. For the CWRU dataset, vibration data collected from both the drive end (CWRU$_d$) and fan end (CWRU$_f$) are used. The fault types considered are Rolling Element (RE), Outer Race (OR), Inner Race (IR), and the Healthy condition. To ensure uniformity, vibration signals from both datasets are resampled to 12 kHz and 20 samples are used for each fault type. The tasks defined for evaluation are summarized in Tab. 1.

**Table 1**. Description of diagnostic tasks for same-domain and cross-domain fault classification.

| Task | Type | Source | Target |
|---|---|---|---|
| T1 | Same-domain | CWRU$_d$ | CWRU$_f$ |
| T2 | Same-domain | CWRU$_f$ | CWRU$_d$ |
| T3 | Cross-domain | CWRU$_d$ | IMS |
| T4 | Cross-domain | CWRU$_f$ | IMS |
| T5 | Cross-domain | IMS | CWRU$_d$ |
| T6 | Cross-domain | IMS | CWRU$_f$ |

### 3.2 Comparison

The classification accuracy of the proposed method is compared with three existing methods: Source HypOthesis Transfer (SHOT) [9], Contrastive Test-Time Adaptation (AdaContrast) [10] and Tent [11]. The accuracy results for each task are presented in Tab. 2.

### 3.3 Observations

The observation of the confusion matrices reveals that SHOT, AdaContrast, and the proposed method successfully cluster the fault types, whereas Tent struggles significantly. However, the relatively low accuracy scores across all methods highlight the inherent challenge of unsupervised domain adaptation in a source-free setting. The slight improvement in accuracy observed with the proposed method can be attributed to the adaptive weight

**Table 2**. Classification accuracy (%) of different methods for same-domain and cross-domain diagnostic tasks.

| Task | SHOT | AdaContrast | Tent | Proposed |
|---|---|---|---|---|
| T1 | 50 | 50 | 22.45 | 50 |
| T2 | 50 | 46.06 | 27.98 | 50 |
| T3 | 50 | 25 | 23.94 | 50 |
| T4 | 0 | 25 | 27.02 | 47.34 |
| T5 | 75 | 25 | 28.94 | 50 |
| T6 | 25 | 25 | 22.13 | 25 |
| Avg | 41.67 | 32.67 | 25.41 | 45.39 |

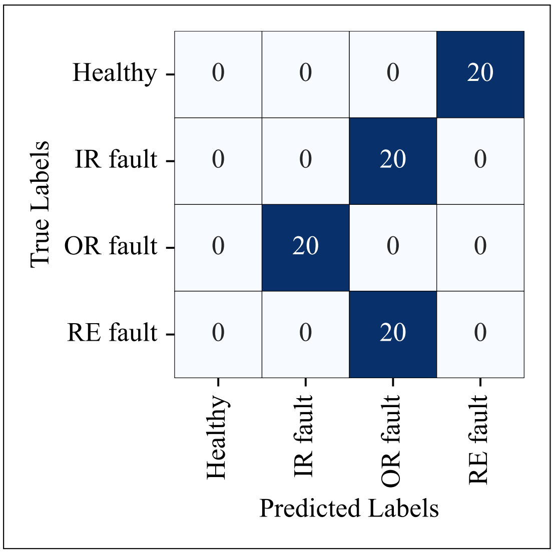


**Figure 2**. Confusion matrix for Task 4 using the SHOT approach.

update strategy, wherein the SOM$_T$ cluster centroids are adjusted towards the nearest cluster centroid in SOM$_S$. This adjustment enhances feature alignment between the source and target domains, leading to better classification performance. For instance, in Task 4 (cross-domain adaptation), the SHOT approach successfully identified three distinct fault types in the target dataset. However, it failed to assign the correct labels, leading to significant misclassification. The corresponding confusion matrix for Task 4 using SHOT is shown in Fig. 2. The proposed method consistently achieved around 50% accuracy in all tasks except Task 6 and was outperformed only in Task 5 by the SHOT approach.

## 4. CONCLUSION

The current study presents a source-free domain adaptation approach for bearing fault diagnosis, leveraging AR modelling, SOM, and a DNN. While AR modelling captures a unique representation of vibration data, the structured feature representation via SOM enhances clustering, leading to improved classification accuracy. Experimental results demonstrate that the proposed method outperforms existing approaches in both same-domain and cross-domain tasks. By addressing the challenge of domain shift in a source-free framework, the proposed method improves the feasibility of deploying machine learning models in real-world fault diagnosis applications. While the results are promising, further research is needed to refine pseudo-labelling strategies for better accuracy and to enhance robustness in noisy environments.

## 5. ACKNOWLEDGMENTS

The authors genuinely thank Shahis Hashim for his constructive feedback and unwavering support throughout the presented research.

## 6. REFERENCES

[1] L. Cui, Z. Jiang, D. Liu, and D. Zhen, “A novel weighted sparse classification framework with extended discriminative dictionary for data-driven bearing fault diagnosis,” *Mechanical Systems and Signal Processing*, vol. 222, p. 111777, 2025.

[2] S. Hashim, S. K. Mishra, and P. Shakya, “An approach for adaptive filtering with reinforcement learning for multi-sensor fusion in condition monitoring of gearboxes,” *Computers in Industry*, vol. 164, p. 104214, 2025.

[3] Y. Wang, F. Jia, J. Shen, and L. Hao, “Source-free domain adaptation network for rolling bearing fault diagnosis,” in *2023 IEEE International Conference on Mechatronics and Automation (ICMA)*, pp. 1691–1696, IEEE, 2023.

[4] H.-W. Yeh, B. Yang, P. C. Yuen, and T. Harada, “Sofa: Source-data-free feature alignment for unsupervised domain adaptation,” in *Proceedings of the IEEE/CVF*

*Winter Conference on Applications of Computer Vision*, pp. 474–483, 2021.

[5] A. K. Tangirala, *Principles of system identification: theory and practice*. CRC press, 2018.

[6] T. Kohonen, "Self-organized formation of topologically correct feature maps," *Biological cybernetics*, vol. 43, no. 1, pp. 59–69, 1982.

[7] C. W. R. U. B. D. Center, "Case western reserve university bearing data center website," 2009.

[8] H. Qiu, J. Lee, J. Lin, and G. Yu, "Wavelet filter-based weak signature detection method and its application on rolling element bearing prognostics," *Journal of sound and vibration*, vol. 289, no. 4-5, pp. 1066–1090, 2006.

[9] J. Liang, D. Hu, and J. Feng, "Do we really need to access the source data? source hypothesis transfer for unsupervised domain adaptation," in *International conference on machine learning*, pp. 6028–6039, PMLR, 2020.

[10] D. Chen, D. Wang, T. Darrell, and S. Ebrahimi, "Contrastive test-time adaptation," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 295–305, 2022.

[11] D. Wang, E. Shelhamer, S. Liu, B. Olshausen, and T. Darrell, "Tent: Fully test-time adaptation by entropy minimization," *arXiv preprint arXiv:2006.10726*, 2020.