Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Automated Hyperparameter Optimization for Transfer Learning-Based Machine Fault Diagnosis Using Hybrid Genetic–Simulated Annealing Algorithm

Mohammed Jarbou

mhjarbou@gmail.com

SUNY Binghamton: Binghamton University https://orcid.org/0000-0002-5350-7898

Daehan Won

SUNY Binghamton: Binghamton University

Sanwon Yoon

SUNY Binghamton: Binghamton University

Research Article

Keywords: Machine fault diagnosis, Predictive maintenance, Transfer learning, Hyperparameter optimization, Genetic simulated annealing

Posted Date: August 4th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-10511415/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Automated Hyperparameter Optimization for Transfer Learning-Based Machine Fault Diagnosis Using Hybrid Genetic–Simulated Annealing Algorithm

Mohammed Jarbou[1*†], Daehan Won[1*†] and Sangwon Yoon[1†]

[1*]School of Systems Science and Industrial Engineering, Binghamton University, 4400 Vestal Pkwy E, Binghamton, 13902, NY, USA.

*Corresponding author(s). E-mail(s): mhjarbou@gmail.com; dhwon@binghamton.edu;
Contributing authors: yoons@binghamton.edu;
†These authors contributed equally to this work.

Abstract

Transfer learning (TL) has attracted growing attention for addressing the performance degradation of Machine Failure Diagnosis (MFD) models in real industrial environments, where challenges such as domain shift and data scarcity are common. However, TL performance is sensitive to hyperparameter configurations and source-target alignment. The existing literature lacks automated, objective mechanisms to optimize TL hyperparameters for MFD. This research addresses this gap by introducing a hyperparameter optimization framework for TL-based MFD using a Genetic Simulated Annealing (GSA) algorithm. GSA is adopted for its balanced exploitation-exploration search capabilities to ensure robust, optimal TL hyperparameters. The framework has been stress-tested across various scenarios, including same-machine and cross-machine failure diagnosis setups, with varying levels of data scarcity and different machine failure types, to ensure its robustness in real-world industrial environments. Experiments using two benchmark datasets (CWRU and SEU) show that the proposed mechanism consistently improves diagnostic performance compared with the baseline TL-based MFD strategy, especially under data scarcity and domain shift scenarios.

Keywords: Machine fault diagnosis, Predictive maintenance, Transfer learning, Hyperparameter optimization, Genetic simulated annealing

1 Introduction

Rotary machines, representing 90% of industrial machinery, are heavily deployed to perform multiple industrial tasks across various industrial sectors [1]. Evolving operating conditions, prolonged operations, and harsh manufacturing environments cause rotary machines to wear out and exhibit various failure modes over time. Early diagnosis of these failures improves operational efficiency and reduces maintenance costs [1]. Data-driven diagnostic approaches have been widely used in the past decade due to their accuracy and reduced reliance on human experience and knowledge [2]. Collecting sufficient labeled samples with identical training and testing distributions is labor-intensive and likely infeasible given the complex and evolving operating conditions in modern manufacturing environments, which challenge the accuracy of data-driven diagnostic methods [2].

That said, growing attention has been dedicated to TL-based methods to tackle the challenges of data scarcity and domain shift. TL relies on the simple idea of leveraging prior knowledge to learn different yet relevant tasks. Many research endeavors employ TL to tackle the data scarcity challenge by using a pre-trained model from a different domain and fine-tuning it to classify machine failures with minimal labeled data [2]. Other research studies focus on the domain shift challenge that arises from unseen fault modes, the deployment of different or new machine(s), or shifts in feature distributions over time [3].

Based on a recent review, the gaps of the TL-based fault diagnosis literature are manifold, including the lack of 1) robust mechanisms to mitigate negative transfer, 2) federated TL frameworks for better generalization, 3) reliable mechanisms to mitigate label noise effect, 4) online mechanisms that adapt in real time, 5) Objective mechanisms for automated TL optimization by selecting the optimal hyperparameters for pre-trained models [4]. This research emphasizes the development of an automated mechanism for fine-tuning the pre-trained model to ensure optimal diagnosis accuracy. The proposed mechanism adopts the GSA algorithm for selecting the optimal transfer learning hyperparameters. The contribution of this study is twofold:

1. Addressing the Failure Diagnosis (FD) literature gap concerning the absence of automated and objective mechanisms for TL optimization in machine FD. To the best of our knowledge, this study is the first to introduce an optimization-based mechanism for selecting TL hyperparameters, including 1) the number of unfrozen layers, 2) the number of retraining epochs, and 3) the number of initial training epochs for the new fully connected layer.
2. Addressing two gaps in TL literature by exploring: 1) the effectiveness of metaheuristic methods for TL optimization, and 2) additional TL hyperparameters that require optimization.

The literature will be further discussed in Section 2 to highlight the contribution of this research. Section 3 will be dedicated to the proposed methodology. The validation procedures and the results will be thoroughly discussed in Sections 4 and 5, respectively. Section 6 will wrap up this study with the conclusion.

2 Literature Review

This section starts with a review of the four main TL methods. Then, further discussion will be dedicated to the parameter-based TL method as the focus of this study. Finally, a review of hyperparameter optimization of TL models will also be provided. Recent studies reviewed the literature of TL-based methods in terms of different perspectives, such as fault feature extraction, industrial applications insights, and the categorization of TL methods [3]. According to [1], TL methods are generally classified into feature-based, instance-based, parameter-based, and adversarial-based methods based on how the knowledge is transferred between the source and target domains. Each of these methods will be discussed in this section, focusing on their applications in machine FD.

2.1 Feature-based TL

Feature-based methods assume common feature patterns across the source and target domains. Thus, these methods create a common feature space that is used to learn the fault features in the target domain. In this context, a cross-device deep TL framework has been developed by [5] for diagnosing bearing faults, leveraging multi-domain information fusion, and employing multi-kernel maximum mean discrepancy. The high computational costs, in addition to the need for human expertise, are the main drawbacks of this method.

2.2 Instance-based TL

These methods handle the situations where the source and target similarity is rooted at the instance levels. The goal of these methods is to learn from the relevant samples in the source domain and mitigate the impact of irrelevant samples. The applications of these techniques encompass the diagnosis of failures in motors, pumps, turbines, and turbochargers [1]. The accuracy of these methods relies on mitigating the negative transfer impact, which requires technical expertise to ensure sampling and transferring the relevant samples only.

2.3 Adversarial-based TL

While the previous two methods assume feature-level and instance-level similarities between the source and target domains, adversarial TL methods handle the challenge of data scarcity and significant discrepancies between the source and target domains by generating synthetic samples through generative adversarial networks. Extensive research work has been done in this regard to generate reliable labeled samples for fault diagnosis of motors, bearings, gearboxes, and pumps. Considering the scope of this study, we refer the readers to [1] for an extensive and detailed review of the applications of adversarial TL for machine FD.

2.4 Parameter-based TL

Parameter-based, also known as model-based, methods leverage the knowledge stored in the parameters of pre-trained models to learn specific tasks in the target domain. The process of knowledge transfer starts with pretraining a model on different yet relevant tasks. The pre-trained model is then fine-tuned to learn different yet relevant tasks in the target domain. The fine-tuning step is accomplished by first freezing the backbone of the model, where general features are learned, then training the model's head, where task-specific features need to be learned. More details will be provided in the methodology section in this regard. CNN-based models are widely used in the TL-based FD literature due to their powerful feature extraction capabilities from image data. Models such as AlexNet [6], VGGNet [7], ResNet [8], Xception [9], and others have proven effective for diagnosing the faults of bearings, gearboxes, pumps, and turbines. The performance of parameter-based methods depends on 1) the similarity between the source and target domains, 2) the availability of labeled data in the target domain, and 3) the values of the transfer learning hyperparameters [1]. Numerous studies have thoroughly investigated the first two issues employing various methodologies, including the alignment of source-domain distributions, the generation of synthetic data to mitigate data scarcity, and the utilization of data-free training techniques that enable the retraining of pretrained models with limited data samples[2]. A recent systematic review [1] highlighted the lack of automated, objective methods for optimizing TL hyperparameters. To the best of our knowledge, this research is the first to introduce an automated, optimization-driven mechanism for fine-tuning the following TL hyperparameters:

1. The amount of knowledge that needs to be transferred: measured by the number of layers to be unfrozen and fine-tuned for the target task.
2. The amount of fine-tuning measured by the number of training epochs to fine-tune the unfrozen layers.
3. The number of training epochs to initially train the new fully connected layers (classifier).

2.5 TL hyperparameter tuning for CNN models

There has been growing interest recently in optimizing the hyperparameters of CNN-based models [10]. Finding the optimal hyperparameters not only improves the model's accuracy but also reduces training time. Different optimization methods, such as Bayesian, metaheuristic, and numerical-based, have been proposed to improve the performance of CNN models by finding the optimal hyperparameters [10]. In this regard, the impact of optimizing different CNN hyperparameters, such as optimizers, learning rate, kernel size, and others, has also been investigated. The optimization of TL hyperparameters remains an area requiring further scholarly investigation. A recent study examined the influence of four optimization techniques on the fine-tuning of pretrained CNN classification accuracy via TL [11]. The findings indicated that hyperparameter tuning in TL significantly enhances CNN classification performance. Similarly, the adjustment of the number of layers subjected to fine-tuning and the learning rate has proven effective in augmenting the accuracy of pretrained CNN models. The study also recommended investigating the performance of other optimization methods, such as genetic-based metaheuristics, as well as exploring other TL hyperparameters. To the best of our knowledge, this recommendation is not addressed specifically in the predictive maintenance and FD literature.

That said, this study builds upon the recommendations of two systematic reviews [1, 11] on automating TL optimization using metaheuristic approaches. Moreover, a comprehensive experimental analysis was conducted to evaluate the proposed methodology's validity, considering factors such as data scarcity, diagnosing faults of different machine types (bearings and gearboxes), and the homogeneity between the source and target domains (same-machine and cross-machine FD).

3 Methodology

This study introduces an end-to-end framework for automated TL-based machine FD. It addresses situations in which diagnostic knowledge from a pretrained model is used to diagnose 1) new failures exhibited by the same machine or 2) failures exhibited by a new machine of a different type. The framework has two phases. During the offline phase, a CNN model is trained with spectrograms of mechanical vibration signals. Fourier analysis creates time-frequency representations of the raw vibration data to enrich the dataset. In the online phase, the model is updated to diagnose a new type of failure using parameter-based TL. The typical update process consists of three steps, namely 1) classifier change, 2) weights initialization, and 3) partial training. In the first step, the classifier is changed to match the new number of failure labels. The initialization step partially trains the weights of the last fully connected layer(s) to enhance the overall update process. The final step freezes part of the pre-trained model and retrains the remaining part to fit the scope of the new diagnostic task. Figure 1 is a schematic illustration of the proposed framework. The following sections will thoroughly discuss the two phases of the proposed framework.

3.1 Offline Phase

3.1.1 Data acquisition

The proposed model begins with collecting Machine Sensory Data (MSD) to train the offline diagnosis model. We used the CWRU and SEU benchmark datasets, common in machinery FD. CWRU and SEU datasets include vibration signals for various bearing and gearbox health conditions. These conditions are monitored under different operating and severity levels, shown in Table 1. The datasets' conditions will validate the framework's effectiveness by mimicking different failure domain shift scenarios, as explained later.

Table 1 Description of the datasets used

CWRU Dataset		
Health condition	Motor load (hp)	Fault diameter (in)
Normal	0	0
Inner race	0	0.021
SE Dataset		
Health condition	Load configuration	
Healthy gear	20 Hz – 0 V	
Chipped tooth	20 Hz – 0 V	

3.1.2 Data processing

Industrial machines operate in noisy, dynamic environments, making MSD characteristics often nonlinear and non-stationary. Traditional feature extraction requires substantial domain expertise and risks information loss [12]. Extracting discriminative high-level features based only on time and frequency data is challenging [13]. Therefore, time-frequency analysis techniques excel in the FD literature for capturing salient features from nonlinear, non-stationary signals [14]. This study uses the FFT method for its simplicity and popularity in the PM domain.

3.1.3 Offline model training

In the offline phase, the CNN model is trained using the spectrograms generated from the last step as input. CNN is widely used for failure diagnosis, considering its superiority in extracting discriminative high-level representations of different MHCs. The general structure of a CNN consists of multiple feature extractor layers as well as a classifier. Features are extracted by passing the input data through a sequence of convolutional and pooling layers. Each convolutional layer has C channels, each of which generates feature maps of size c. The feature map $x_c^p \in R^{a \times a}$ of channel $c \in C$ in layer p is calculated using the feature map $x_c^{p-1} \in R^{\hat{a} \times \hat{a}}$ of the previous layer using a non-linear activation function A as illustrated in

Equation (1).

$$x_c^p = A\left(x_c^{(p-1)} * \mathbf{W}_c^p + \mathbf{B}_c^p\right) \quad (1)$$

Where $\mathbf{W}_c^p \in \mathbb{R}^{m\times n}$ denotes the weight matrix associated with channel c of layer p, and $\mathbf{B}_c^p \in \mathbb{R}$ represents the corresponding bias term of the c^{th} channel in layer p. The function $A(\cdot)$ denotes the adopted activation function, while $*$ refers to the convolutional operator.

To avoid overfitting, the spatial dimensions of the generated feature maps are reduced using pooling layers, which help extract higher-level feature representations and improve the generalization capability of the trained model, as described in Equation (2). In this context, down$(\cdot)$ denotes the subsampling (pooling) operation applied to the input feature maps, without introducing learnable parameters.

$$\mathbf{x}_c^l = A\left(\mathbf{W}_c^l * \text{down}\left(\mathbf{x}_c^{(l-1)}\right) + \mathbf{B}_c^l\right) \quad (2)$$

The training mechanism of the CNN model follows the standard backpropagation algorithm to minimize the total loss, represented by the categorical cross-entropy function shown in Equation (3). In this formulation, a_{bj} and $\hat{a}_{bj}$ denote the true and predicted probability of class b for sample j, respectively.

$$\mathcal{L}_{\text{CNN}} = -\sum_{j=1}^{M}\sum_{b=1}^{B} a_{bj}\log\left(\hat{a}_{bj}\right) \quad (3)$$

3.2 Online phase

In the online phase, we assume that either the machines exhibit novel types of failures as they run or a pretrained diagnostic model needs to be updated to diagnose failures in a new machine of a different type. To address these scenarios, the proposed framework includes an online mechanism for incrementally updating the pre-trained model to learn new failures. More details will be provided in the following sections.

3.2.1 Model updating

The pre-trained CNN model remains valid for diagnosing the incoming samples as long as the new samples follow the same distribution as those in the trained data. The occurrence of new failure(s) changes the distribution of the new samples such that $P(Y) \neq P\left(\grave{Y}\right)$, where Y and $\grave{Y}$ are the labels of the training and testing samples, respectively. In this situation, TL is commonly used to incrementally update the pre-trained CNN model to learn the new failure types. In this study, parameter-based TL is employed to update the pre-trained model following three steps. First, the classifier is changed to fit the number of the new failure types. Then, the weights of the fully connected layers are initialized and partially trained to fit the new learning context. In the last step, part of the feature extraction layers, in addition to the fully-connected layers, are fine-tuned. The updating process is illustrated in Figure 2.

3.2.2 TL hyperparameters optimization

The performance of the TL-based updating mechanism depends on the value of different transfer parameters, which are 1) the updating depth, 2) the amount of updating, and 3) the amount of initial tuning. The updating depth refers to the number of layers that need to be updated. Similarly, the amount of updating is the number of training epochs required to learn the discriminative characteristics of the new failure type(s). The amount of initial training is the number of training epochs to initially train the fully connected layers whose weights are initialized with random values. No study offers a comprehensive framework for updating pre-trained models with optimal transfer parameters. This study fills that gap using a hybrid GSA algorithm, which overcomes the limitations of Simulated Annealing (SA) and Genetic Algorithm (GA) [15]. GA is good at global search but weak in local search near convergence, while SA excels locally but less globally [15]. Combining them improves performance, as shown in previous research [15]. GSA is a well-established algorithm, so the complete description of the GSA steps is out of the scope of this study.

4 Validation Experiments

The proposed framework is validated across two FD scenarios, wherein the machines in the target

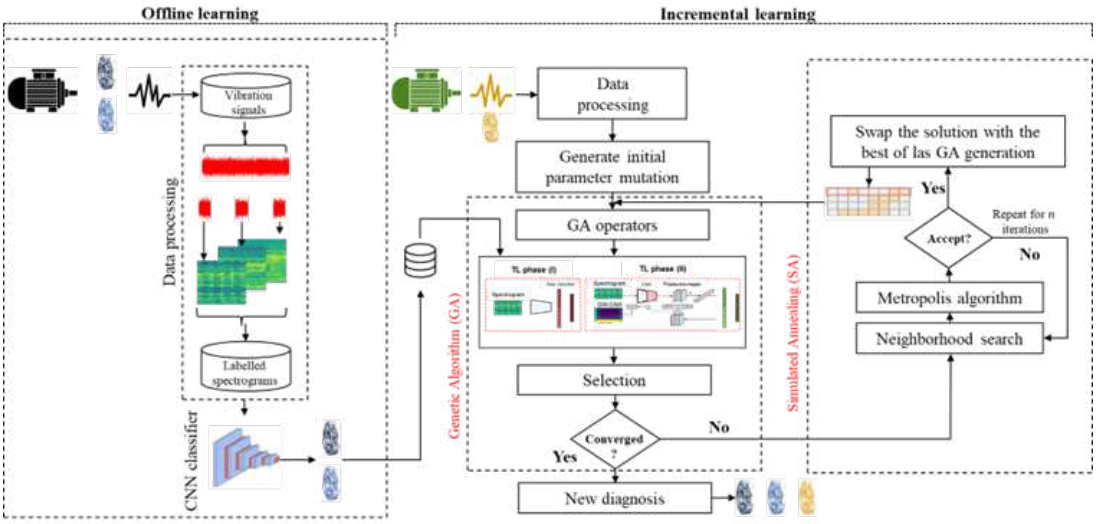


Fig. 1 Schematic illustration of the proposed framework

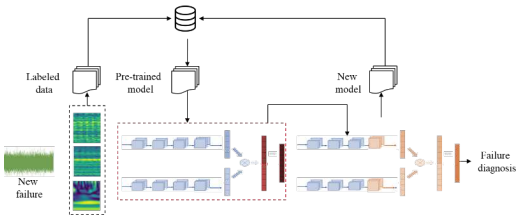


Fig. 2 Model updating mechanism

and source domains are either the same or different. In each scenario, the framework is validated for bearing and gearbox failures. For scenarios involving similar source-target domains, the validation is conducted through four tasks for each of the bearing and gearbox, separately. These examples are summarized in Table 2. Similarly, four different source-target domains, the framework is validated by considering four tasks for diagnosing gearbox faults employing a pretrained bearing model, and four tasks for diagnosing bearing faults utilizing a gearbox-trained model. These examples are summarized in Table 3.

The framework is also validated by considering the impact of data scarcity, measured by how input vibration signals are processed using the sliding window technique. The length of the sliding window and the amount of overlap determine the number and quality of the spectrograms produced from the same input. For instance, smaller windows combined with a high overlap yield more and higher-quality spectrograms. The framework is tested under four data scarcity scenarios, defined by the values of window length and overlap amount, as summarized in Table 4.

Table 2 Same-machine fault diagnosis validation scenarios.

CWRU Dataset		
Task	**Pretrained model**	**New failure**
T1	Normal–IR007–B021	B007
T2	Normal–B021–B007	IR007
T3	Normal–B007–IR007	OR021@12
T4	Normal–IR021–B021	OR021@3
SEU Dataset		
Task	**Pretrained model**	**New failure**
T1	Normal–Surface–Root	Missing tooth
T2	Normal–Miss–Chipped	Surface
T3	Normal–Chipped–Surface	Tooth root
T4	Normal–Root–Miss	Chipped tooth

The performance of the proposed framework is also compared to two common tuning mechanisms, namely full and partial training of the pretrained model. In this regard, the mean and standard deviation of the model's accuracy, sensitivity, and specificity are calculated by repeating each experiment 30 times to account for the randomness of the adopted CNN-based framework. To show the significance of optimizing the TL hyperparameters, the metrics mean and standard deviation of the proposed mechanism are statistically compared to those of full training, as well

Table 3 Different-machine fault diagnosis validation scenarios.

CWRU → SEU		
Task	**Pretrained model**	**New failure**
T1	Normal–IR007–B021	Root tooth
T2	Normal–B021–B007	Chipped tooth
T3	Normal–IR021–B021	Surface tooth
T4	Normal–IR021–B021	Missing tooth
SEU → CWRU		
Task	**Pretrained model**	**New failure**
T1	Normal–Surface–Root	B007
T2	Normal–Miss–Chipped	IR007
T3	Normal–Root–Miss	OR021@3
T4	Normal–Chipped–Surface	OR021@12

Table 4 Data scarcity controlling parameters.

Window length	**Signal overlap**
1024	No overlap
1024	0.5
2048	No overlap
2048	0.5

as partial training mechanisms using $t-test$. The complete experimental design conducted for the framework validation is summarized in Table 5.

Table 5 Factor level combinations for experimental design.

Factor	**Level 1**	**Level 2**
Domain similarity	Same-machine	Cross-machine
Window length	1024	2048
Window overlap	No overlap	0.5

5 Results and Analysis

This section provides an in-depth analysis of the validation experiments discussed in the previous section. The analysis is split into two parts, covering scenarios with the same and different source-target domains. The procedures aim to assess the effect of optimizing TL hyperparameters on the model's: 1) performance, evaluated by specificity, in comparison with two methods where TL is not optimized; and 2) robustness under data scarcity.

5.1 Scenario I: Same source-target domains

In the first scenario, the framework will be validated to diagnose new machine failures that belong to the same machine. The CWRU and SEU datasets will be used to validate the proposed framework for diagnosing unseen bearing and gearbox failures. Four experiments, representing different TL tasks, are conducted for each dataset to diagnose four unseen failures for each machine type, as summarized in Table 2. Figure 4 and Figure 5 demonstrate that optimizing the TL hyperparameters significantly improves the model's testing specificity across all four TL tasks for both machine types. The improvement in specificity is statistically significant, with a $p-value \leq 0.05$, based on 30 repeated experimental runs conducted to account for the randomness inherent in CNN training. Diagnosing unseen failures presents considerable challenges due to the limited availability of labeled samples. This study addresses this issue by verifying the efficacy of optimizing TL hyperparameters in scenarios where labeled samples of unseen failures are scarce. In this study, data scarcity is simulated by assuming that only 50% of the available spectrograms are available for the unseen failure types. Figure 6 and Figure 7 show that optimizing the TL hyperparameters significantly improves testing specificity under conditions of data scarcity. The improvement in specificity is statistically significant, with a $p-value \leq 0.05$, based on 30 repeated experimental runs conducted to account for the randomness inherent in CNN training.

5.2 Scenario II: different source-target domains

This section repeats the validation process from scenario I to diagnose machine failures across various machine types. In this scenario, we assume that a pretrained diagnosis model is used to diagnose failures of a new machine of a different type.

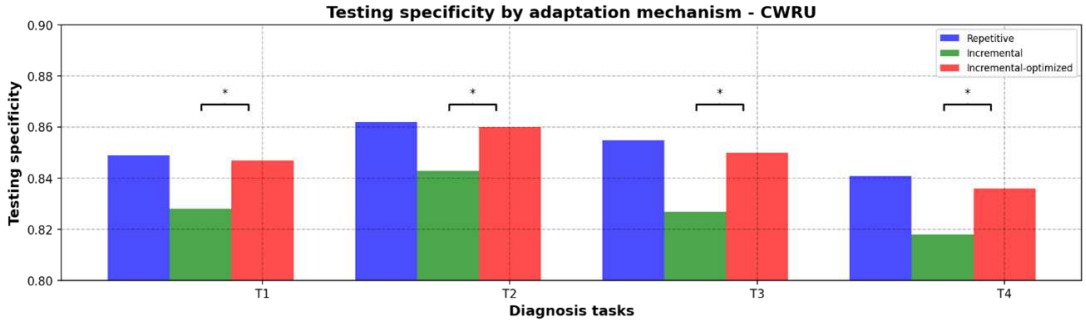


Fig. 3 Optimized vs. unoptimized specificity by bearings FD tasks (Scenario I: CWRU)

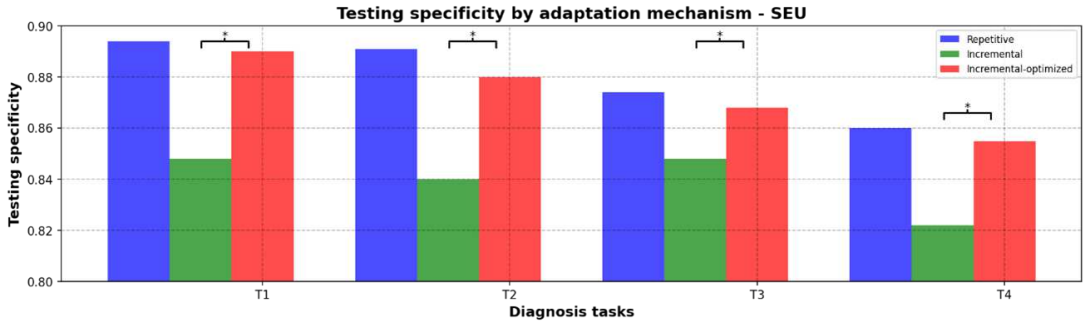


Fig. 4 Optimized vs. unoptimized specificity by gearbox FD tasks (Scenario I: SEU)

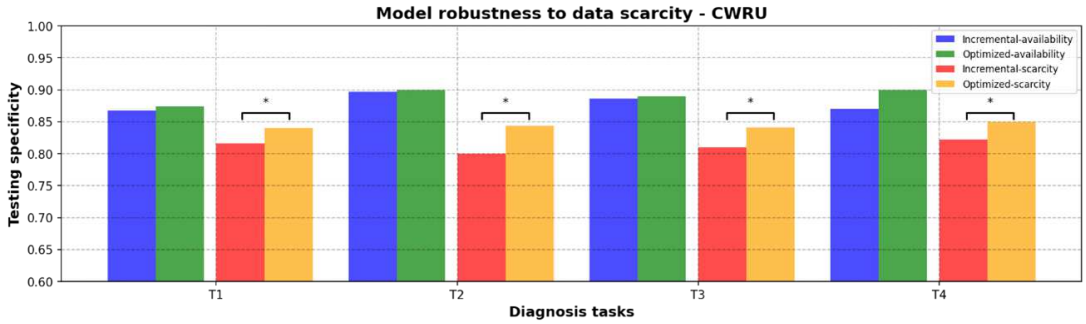


Fig. 5 Optimized vs. unoptimized specificity by bearings FD tasks under data scarcity (Scenario I: CWRU)

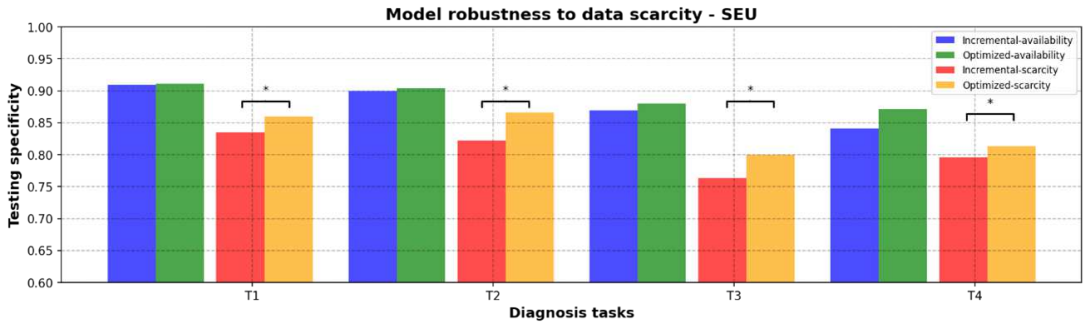


Fig. 6 Optimized vs. unoptimized specificity by gearbox FD tasks under data scarcity (Scenario I: SEU)

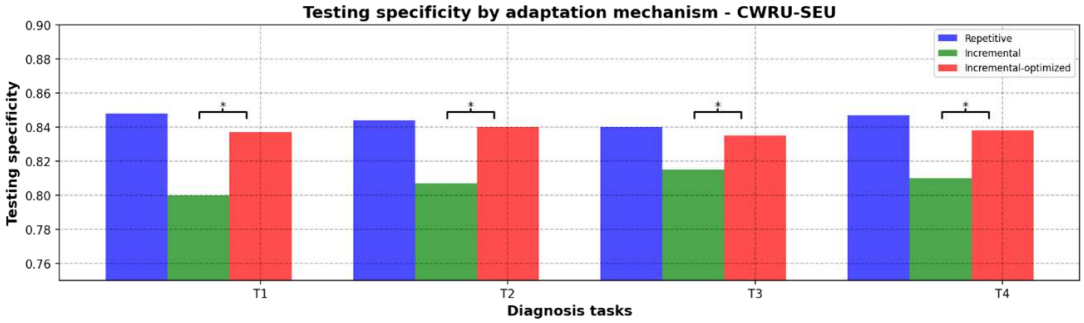


Fig. 7 Optimized vs. unoptimized specificity by FD tasks (Scenario II: CWRU → SEU)

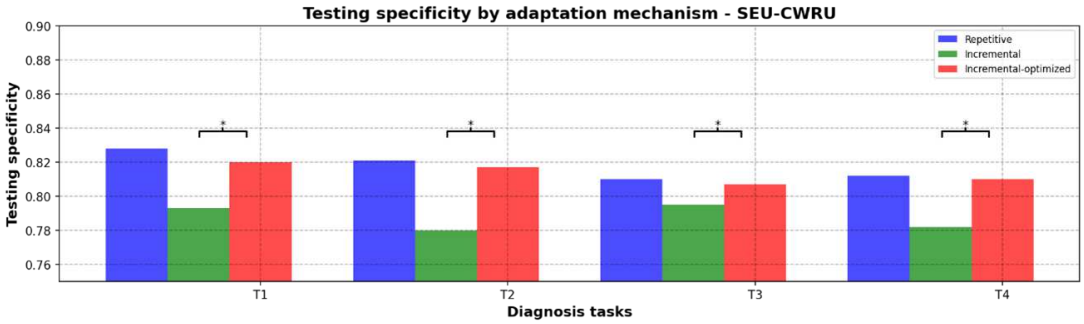


Fig. 8 Optimized vs. unoptimized specificity by FD tasks (Scenario II: SEU → CWRU)

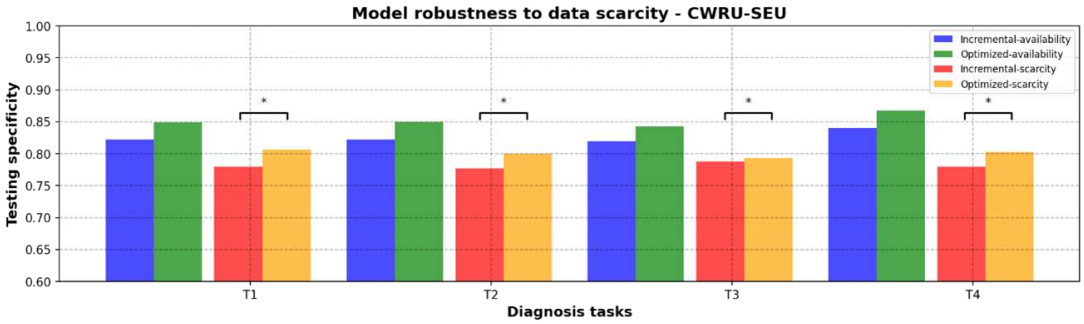


Fig. 9 Optimized vs. unoptimized specificity by FD tasks under data scarcity (Scenario II: CWRU → SEU)

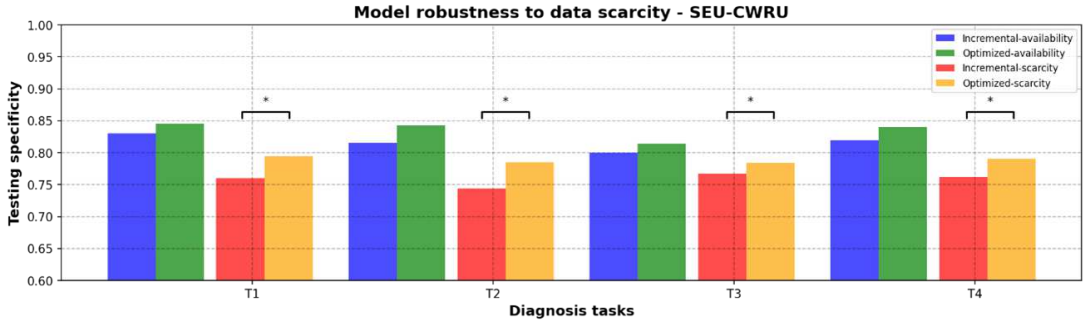


Fig. 10 Optimized vs. unoptimized specificity by FD tasks under data scarcity (Scenario II: SEU → CWRU)

Two experimental sets are performed, each with four distinct TL tasks, to diagnose gearbox failures using knowledge from the bearings pre-trained diagnosis model and vice versa. Table 3 lists the TL tasks used in this scenario.Similarly to the previous section, the CWRU and SEU datasets are used to fulfill the validation scope in this section. The same metrics of accuracy and robustness to data scarcity are also adopted to validate the effectiveness of optimizing TL hyperparameters for this scenario. Figure 8 through Figure 10 confirm the results of Scenario I, showing that optimizing TL significantly improves diagnostic performance across most of the tasks, with a $p - value \leq 0.05$, based on 30 repeated experimental runs conducted to account for the randomness inherent in CNN training.

6 Conclusion

The study proposes an automated mechanism to optimize TL hyperparameters for diagnosing unseen failure types in a time-efficient manner. The GSA algorithm is proposed for TL hyperparameter optimization as a powerful optimization technique. The effectiveness of the proposed GSA mechanism for automating TL optimization has been validated through extensive validation procedures that address: 1) two types of machinery (bearings and gearboxes), 2) two scenarios of knowledge transfer, wherein the source and target domains may pertain to identical or different machines, and 3) data scarcity, thereby emulating real operational environments where labeled samples are limited. Furthermore, each experiment was conducted 30 times to mitigate the variability introduced by the stochastic nature of CNN training. The findings demonstrate that optimizing TL hyperparameters using GSA significantly improves diagnostic specificity, considering the aforementioned experimental conditions. As this study addresses the existing gap in automated and objective TL optimization mechanisms within the literature, future investigations could explore alternative optimization algorithms and more complex TL tasks, such as multi-source TL, federated TL, and multi-task TL. Future research could also explore other TL hyperparameters to understand the impact of each parameter.

Declarations

The authors declare that no funds, grants, or other support were received during the preparation of this manuscript.

- Author contribution
  Mohammed Jarbou: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Methodology, Formal analysis, Conceptualization. **Daehan Won**: Writing – review & editing, Supervision, Conceptualization, Validation. **Sangwon Yoon**: Conceptualization, Supervision

- Funding
  Not applicable

- Conflict of interest/Competing interests (check journal-specific guidelines for which heading to use)
  The authors have no relevant financial or non-financial interests to disclose

- Ethics approval and consent to participate
  Not applicable

- Consent for publication
  Not applicable
- Data availability
  Not applicable
- Materials availability
  Not applicable
- Code availability
  Not applicable

References

[1] Feisa, T.T., Gebremedhen, H.S., Kibrete, F., Woldemichael, D.E., Ageze, S.D., Nega, H.D.: A comprehensive review of deep transfer learning for intelligent fault diagnosis in rotating machines. Engineering Reports **7**(11), 70494 (2025) https://doi.org/10.1002/eng2.70494

[2] Chen, X., Yang, R., Xue, Y., Huang, M., Ferrero, R., Wang, Z.: Deep transfer learning for bearing fault diagnosis: A systematic review since 2016. IEEE Transactions on Instrumentation and Measurement **72**, 1–21 (2023) https://doi.org/10.1109/TIM.2023.3244237

[3] Wang, C., Wang, Z., Liu, Q., Dong, H., Liu, W., Liu, X.: A comprehensive survey on domain adaptation for intelligent fault diagnosis. Knowledge-Based Systems, 114109 (2025) https://doi.org/10.1016/j.knosys.2025.114109

[4] Ahmad, S., Lavin, A., Purdy, S., Agha, Z.: Unsupervised real-time anomaly detection for streaming data. Neurocomputing **262**, 134–147 (2017) https://doi.org/10.1016/j.neucom.2017.04.070

[5] Li, J., Ye, Z., Gao, J., Meng, Z., Tong, K., Yu, S.: Fault transfer diagnosis of rolling bearings across different devices via multi-domain information fusion and multi-kernel maximum mean discrepancy. Applied Soft Computing **159**, 111620 (2024) https://doi.org/10.1016/j.asoc.2024.111620

[6] Ma, P., Zhang, H., Fan, W., Wang, C., Wen, G., Zhang, X.: A novel bearing fault diagnosis method based on 2d image representation and transfer learning-convolutional neural network. Measurement Science and Technology **30**(5), 055402 (2019) https://doi.org/10.1088/1361-6501/ab0793

[7] Zhou, J., Yang, X., Zhang, L., Shao, S., Bian, G.: Multisignal vgg19 network with transposed convolution for rotating machinery fault diagnosis based on deep transfer learning. Shock and Vibration **2020**(1), 8863388 (2020) https://doi.org/10.1155/2020/8863388

[8] Chen, Z., Cen, J., Xiong, J.: Rolling bearing fault diagnosis using time-frequency analysis and deep transfer convolutional neural network. Ieee Access **8**, 150248–150261 (2020) https://doi.org/10.1109/ACCESS.2020.3016888

[9] Zhang, W., Zhang, T., Cui, G., Pan, Y.: Intelligent machine fault diagnosis using convolutional neural networks and transfer learning. IEEE Access **10**, 50959–50973 (2022) https://doi.org/10.1109/ACCESS.2022.3173444

[10] Raiaan, M.A.K., Sakib, S., Fahad, N.M., Mamun, A.A., Rahman, M.A., Shatabda, S., Mukta, M.S.H.: A systematic review of hyperparameter optimization techniques in convolutional neural networks. Decision Analytics Journal **11**, 100470 (2024) https://doi.org/10.1016/j.dajour.2024.100470

[11] Öcal, A., Koyuncu, H.: An in-depth study to fine-tune the hyperparameters of pre-trained transfer learning models with state-of-the-art optimization methods: Osteoarthritis severity classification with optimized architectures. Swarm and Evolutionary Computation **89**, 101640 (2024) https://doi.org/10.1016/j.swevo.2024.101640

[12] Namuduri, S., Narayanan, B.N., Davuluru, V.S.P., Burton, L., Bhansali, S.: Deep learning methods for sensor based predictive maintenance and future perspectives for electrochemical sensors. Journal of The Electrochemical Society **167**(3), 037552 (2020)

https://doi.org/10.1149/1945-7111/ab67a8

[13] Chen, R., Huang, X., Yang, L., Xu, X., Zhang, X., Zhang, Y.: Intelligent fault diagnosis method of planetary gearboxes based on convolution neural network and discrete wavelet transform. Computers in industry **106**, 48–59 (2019) https://doi.org/10.1016/j.compind.2018.11.003

[14] Tang, S., Yuan, S., Zhu, Y.: Data preprocessing techniques in convolutional neural network based on fault diagnosis towards rotating machinery. Ieee Access **8**, 149487–149496 (2020) https://doi.org/10.1109/ACCESS.2020.3012182

[15] Li, Y., Guo, H., Wang, L., Fu, J.: A hybrid genetic-simulated annealing algorithm for the location-inventory-routing problem considering returns under e-supply chain environment. The Scientific World Journal **2013**(1), 125893 (2013) https://doi.org/10.1155/2013/125893