



OPEN

# Deep transfer learning strategy for efficient domain generalisation in machine fault diagnosis

Supriya Asutkar<sup>1,2</sup> & Siddharth Tallur<sup>2</sup>✉

Automated fault diagnosis algorithms based on vibration sensor recordings play an important role in determining the state of health of the machines. Data-driven approaches demand a large amount of labelled data to build reliable models. The performance of such lab-trained models degrades when deployed in practical use cases in the presence of distinct distribution target domain datasets. In this work, we present a novel deep transfer learning strategy that fine-tunes the trainable parameters of the lower (convolutional) layers with respect to the changing target domain datasets and transfers the parameters of the deeper (dense) layers from the source domain for efficient domain generalisation and fault classification. The performance of this strategy is evaluated by considering two different target domain datasets and studying the sensitivity of fine-tuning individual layers in the networks using time-frequency representations of the vibration signals (scalograms) as inputs. We observe that the proposed transfer learning strategy yields near-perfect accuracy, even for use cases where low-precision sensors are used for data collection and unlabelled run-to-failure data with a limited number of training samples.

Automated fault diagnosis in machines based on sensor data is an emerging area of interest in condition-based monitoring (CBM) and industrial internet of things (IIoT), that enables improved production efficiency and lower risk of accidents in complex mechanical systems<sup>1</sup>. Unlike monitoring parameters such as surface temperature, power consumption, and presence of wear particles in lubricants, vibration-based fault diagnosis is now a well-established method for identifying incipient faults<sup>2</sup>. Methods for fault diagnosis can be divided into two broad categories: physics-based and data-driven approaches. Physics-based models require establishing a relationship between fault signatures in sensor parameters and progression of damage in mechanical parts of the machinery and typically encompass designing models based on dynamics, finite element method, and modal analysis e.g. parametric model for planetary gears developed by Xue et al.<sup>3</sup>. Implementation of such methods requires an in-depth understanding of the dynamics of the machinery, and customisation for every sensor installation, to account for variations in the ambient environment, mounting, and manufacturing parameters. Several data-driven algorithms have been reported for machine health monitoring, ranging from simple classifiers such as support vector machines (SVM)<sup>4–6</sup>, advanced classifiers like least square interactive support matrix machine (LSISMM)<sup>7</sup>, artificial neural network (ANN)<sup>8</sup> to complex approaches based on variations of deep convolutional neural networks (CNNs)<sup>9–13</sup>.

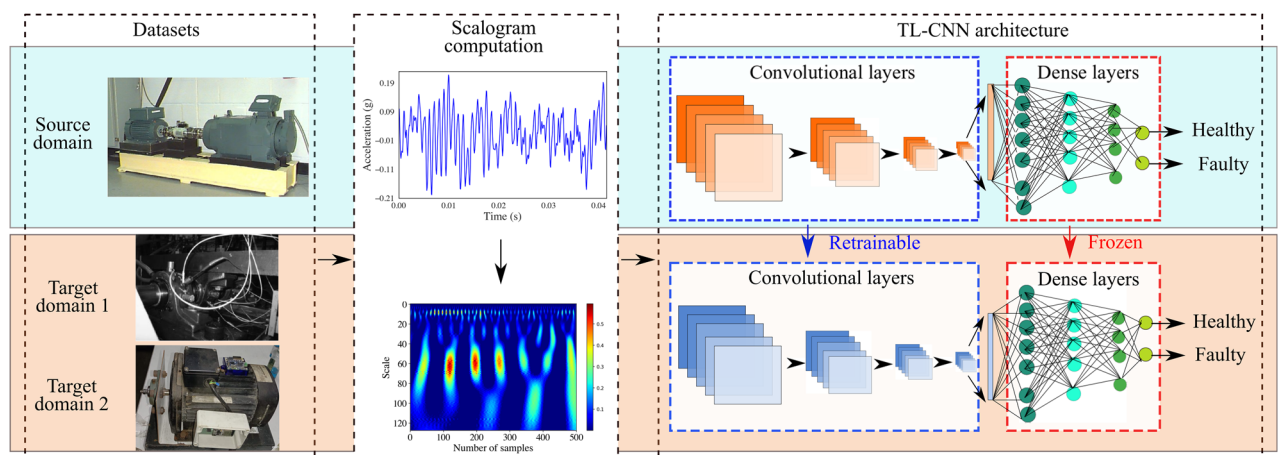
Although significant improvement in fault classification accuracy is possible using deep learning techniques, their performance is contingent on two key requirements: (a) data used for training (source domain) and testing (target domain) should belong to the same distribution, and (b) the necessity of a large labelled training dataset for higher accuracy. This is impractical in real-world scenarios since data distribution is sensitive to changes in ambient environment and installation, variation in power and load, etc. This change in data distribution from the source domain dataset used for training the algorithm to the target domain dataset used for testing is labelled as *domain shift*. To mitigate the problem of domain shift in deep learning models, various transfer learning (TL) strategies are employed wherein knowledge gained from the source domain is leveraged to improve classification ability in the distinctly distributed target domain datasets. Among the various approaches, feature-based TL has shown good domain adaptation capability by reducing the distribution difference between the source and the target domain dataset. Xiao et al. have proposed a domain adaptive motor fault diagnosis technique that uses CNN to extract multi-level features from the raw vibration data and maximum mean discrepancy (MMD) is adopted in the training process to reduce the distribution difference between the source and the target domain

<sup>1</sup>Centre for Research in Nanotechnology & Science (CRNTS), IIT Bombay, Powai, Mumbai, Maharashtra 400076, India. <sup>2</sup>Department of Electrical Engineering (EE), IIT Bombay, Powai, Mumbai, Maharashtra 400076, India. ✉email: stallur@ee.iitb.ac.in

dataset. Thus, the knowledge learned from the source domain that possesses labelled motor vibration data under invariant working conditions is used to improve fault classification accuracy when the target domain belongs to the unlabelled data under constantly varying working conditions<sup>14</sup>. Zhao et al. have proposed an intelligent gearbox fault diagnosis method based on adaptive intraclass and interclass convolutional neural network (AIICNN) to improve generalisation under varying working conditions<sup>15</sup>. In another study, a deep convolution transfer learning network (DCTLN) is proposed that comprises a condition recognition module along with a domain adaptation module to effectively learn domain-invariant features of the target domain datasets. The DCTLN trained with labelled data of the source domain dataset is able to effectively classify unlabelled data of the target domain dataset<sup>16</sup>. Tong et al. have presented an unsupervised fault diagnosis approach incorporating feature TL to efficiently adapt to the varying working conditions of the target domain dataset<sup>17</sup>. Qian et al. on the other hand, have proposed an improved joint distribution adaptation (IJDA) technique to align not only the marginal but conditional distributions of the source and the target domain datasets for effective domain generalisation when validated using bearing and gearbox vibration signal datasets<sup>18</sup>. However, these techniques typically require training architectures from scratch and thus, can be computationally expensive.

Parameter-based TL on the other hand is a commonly employed strategy, wherein only a few layers of large pre-trained networks (e.g. VGG-16, ResNet-50, etc.) are fine-tuned while freezing a large number of the layers, for domain generalisation to a diverse set of target image datasets<sup>19</sup>. This approach is suitable for CBM applications that represent the vibration signals as encoded images through suitable time-frequency transforms e.g. spectrogram, scalogram, etc. Shao et al. have presented a novel deep TL framework comprising a pre-trained VGG-16 network that is trained on the ImageNet dataset which later is fine-tuned using the time-frequency images (scalograms) of the target machine vibration dataset. In the process, only the three highest-level blocks of the pre-trained VGG-16 network are fine-tuned while leaving the weights of the bottom blocks frozen to achieve state-of-the-art accuracy when validated against the three publicly available machine vibration data<sup>20</sup>. Wen et al. proposed a TL strategy that includes the use of a pre-trained ResNet-50 network to identify fault by fine-tuning just the fully connected layer added on top of the ResNet-50 network with respect to the publicly available machine vibration datasets<sup>21</sup>. Since these pre-trained networks are trained on the ImageNet dataset, consisting of images significantly different from the vibration encodings of the machines (spectrograms and scalograms), large networks are required for good adaptation ability. This results in a large number of trainable parameters and memory consumption along with additional post-processing on the target domain images. To address this challenge, recently, several reports for domain adaptation in vibration CBM using TL-CNNs operating on raw time-series vibration data have been reported, albeit with feature transfer and domain adversarial network or parameter transfer in dense layers of CNN<sup>22–24</sup>. However, in most TL reports for CBM, high-precision sensors or publicly available high-resolution datasets are used in the target domain, and domain generalisation of TL models for vastly different target domain data with low-precision sensors has not been adequately explored. CBM with low-cost consumer-grade sensors is essential for scaling the benefits of predictive maintenance to large factories, wherein instrumenting every machine with multiple high precision and expensive piezoelectric vibration sensors is impractical<sup>25,26</sup>. Another challenge in the deployment of deep learning techniques pertains to the acquisition of faulty data sets in the field, as machines may stop working abruptly and therefore result in a disproportionate amount of healthy data being collected over data in faulty operating conditions<sup>27,28</sup>.

In order to overcome the shortcomings above, in this work, we have implemented a modified TL-based classification model wherein a significantly lesser number of layers are used compared to the conventional pre-trained networks to address the challenges of domain shift and data insufficiency. An overview of the TL-CNN architecture proposed in this work is shown in Fig. 1. The work utilises the widely used Case Western Reserve University bearing dataset (CWRU)<sup>29</sup> as the source domain dataset for training CNNs on time-frequency domain representations of the vibration signals, namely scalograms. Whereas, two different target domain datasets are considered as the test dataset to demonstrate the versatility of the proposed TL approach: data generated in the lab from a motor instrumented with low precision sensor (LPS), and run-to-failure unlabelled dataset from the



**Figure 1.** Overview of the TL-CNN architecture presented in this work.

Center for Intelligent Maintenance Systems, University of Cincinnati (IMS)<sup>30</sup>. Measurements obtained with low precision sensors along with high variability (greater interquartile range about the mean) make it difficult to distinguish between distributions for healthy and faulty operating conditions. No additional processing of target domain data is required since the source and target data are subjected to the same time-frequency transform. We have evaluated the performance of deep TL models by transferring parameters of convolutional layers and dense layers and studied the sensitivity of retraining of each layer separately. The key contributions of this paper are as follows:

- Our observations indicate that TL using scalograms as input features yields higher accuracy even when the statistical distribution of data obtained with lower precision sensors is not sufficiently distinct for healthy and faulty classes.
- The novelty of the proposed deep TL algorithm is the improvement in prediction ability obtained with a significantly lesser number of trainable parameters (1320) by fine-tuning (retraining) the convolutional layers (i.e. freezing dense layers) to better capture domain-specific abstract information, as compared to the conventional approach of retraining dense layers (i.e. freezing convolutional layers).
- The improvement in accuracy is also observed for unlabelled data with the limited number of samples available for training (IMS).

## Methodology

**Time-frequency based representations of vibration signals.** Various methods are used to represent vibration signals in the time and frequency domain. However, time-frequency imaging methods are efficient at capturing the non-stationary nature of the vibration signals<sup>31–34</sup>. In this work, we have considered scalograms obtained using continuous wavelet transform (CWT) as a representation of the vibration signals. CWT provides a more flexible time-frequency resolution as compared to spectrograms obtained with short-time Fourier transform (STFT), by using the mother wavelet function that can be scaled and shifted in correspondence to events in the signals. The scalogram is the absolute value of the CWT of a signal, plotted as a function of scale ( $s$ ) and time translation ( $\tau$ ):

$$F(\tau, s) = \frac{1}{\sqrt{|s|}} \int_{-\infty}^{+\infty} f(t) \psi^* \left( \frac{t - \tau}{s} \right) dt, \quad (1)$$

where,  $\psi(t)$  denotes the mother wavelet, and the asterisk denotes the complex conjugate. Expanded wavelets (large values of  $s$ ) capture the low-frequency components of the signal with bad time resolution whereas, shrunken wavelets (small values of  $s$ ) capture high-frequency components with good time resolution.

**Description of experimental datasets.** Brief descriptions of the source and target domain vibration datasets used in this study are provided below:

- CWRU data set* Among various publicly available machine vibration datasets, CWRU is widely accepted as a standard reference labelled dataset that has been used to validate different fault diagnosis models<sup>29</sup>. Vibration signals from a 2 hp Reliance Electric motor were acquired using high-precision accelerometers with healthy and faulty bearings. Single point faults (inner race, ball bearing, outer race) ranging from 7 to 40 mil in diameter were introduced separately in the test bearings through electro-discharge machining. These faulty bearings were reinstalled into the motor at the drive end and fan end. Sensors were placed at these locations and also at the base end of the motor, and vibration data was recorded at a sampling rate of 12 kHz for motor speeds varying from 1797 to 1720 rpm.
- IMS data set* The bearing dataset provided by IMS is available for download from the NASA Ames Prognostics Data Repository<sup>30</sup>. The data set comprises of measurements from high-sensitivity quartz ICP<sup>®</sup> accelerometers, one each installed on four bearings on a shaft. Data from three run-to-failure experiments are provided. The rotation speed was 2000 rpm for a radial load of 6000 lb applied to the shaft. All failures occurred after exceeding 100 million revolutions. Each time series was recorded in an individual file, containing 20,480 data points obtained at 20 kHz sampling rate. Out of the three datasets, we have considered the dataset with the highest number of recordings: channel 3 from test number 3 (culminating in outer race fault). For further details on the experimental setup, please refer report by Qiu et al.<sup>35</sup>. Since the IMS-bearing data set is an unlabelled dataset, the mean peak frequency (MPF) of the spectrogram was used to label the healthy and faulty data, in accordance with the strategy presented by Mukherjee et al.<sup>36</sup>.
- Lab generated data with low precision sensor (LPS)* This dataset comprises vibration data acquired from a motor running at a speed of approximately 3000 rpm using STMicroelectronics STEVAL-STWINK1B wireless industrial sensor node that includes an ultra-wide bandwidth (6 kHz), low-noise, 3-axis vibration sensor (STMicroelectronics IIS3DWB) with 26.7 kHz data rate. The dataset contains 650 files with 15,000 samples (i.e. 560 ms duration) each for both the healthy and faulty operating conditions of the motor. The fault was introduced by coupling a bearing with a faulty ball bearing to the motor shaft.

In our study, CWRU was chosen as the source domain dataset, and IMS and LPS datasets were used as the target domain datasets. The CWRU dataset consists of a vast collection of labelled vibration data obtained with high precision sensor corresponding to baseline (healthy) operation, and various types of bearing faults with

different fault depths obtained with different motor speeds, and is, therefore, a suitable choice for source domain to incorporate more variation and develop a robust deep CNN based fault classification algorithm. While the IMS dataset is also acquired using a high-precision vibration sensor, it is a run-to-failure dataset consisting of a limited number of vibration recordings in faulty operation (useful for studying the TL model performing with insufficient training data). The LPS dataset is altogether different from the CWRU dataset, consisting of data obtained on a different motor running at higher speed, with a low-cost albeit comparatively lower precision MEMS vibration sensor (useful for performance evaluation of domain adaptation of TL model).

**Transfer learning for machine fault diagnosis.** For generating scalograms, the following parameters were used—number of data samples used to generate scalogram: 500, scale factor: 128, wavelet function: Morlet. This resulted in scalogram images of size  $128 \times 500$ . The generated scalograms were normalized using the StandardScaler normalisation method before using them as inputs for the CNN-based classification algorithm. To find the optimum CNN architecture for the training data (CWRU), *RandomSearch* algorithm was employed, wherein the best set of hyperparameters were obtained based on the maximum cross-validation score. Table 1 presents the hyperparameter range used in the *RandomSearch* algorithm and the optimum parameter values obtained from the algorithm that yielded 100% accuracy for fivefold cross-validation with 10 iterations on the training data.

A detailed summary of the CNN architecture realised utilising the tuned hyperparameters is presented in Table 2. The CNN architecture comprises two convolutional layers and three dense layers. Max-pooling layers were used to reduce feature dimensions while retaining salient information of the inputs in the convolutional layers. The convolutional layers incorporated rectified linear unit (ReLU) activation functions to introduce non-linearity, and zero padding was performed to ensure the same input and output image dimensions. The dropout rate of 0.2 was used to avoid over-fitting to the data. The CNN architecture was trained using Adam optimizer with learning rate = 0.01 for 30 epochs using binary cross-entropy as the loss function. We observed that the training accuracy and loss converged within 30 epochs. The predicted probabilities were assessed using a sigmoid function to determine the class to which the input data belongs.

Parameter-based TL allows the reuse of parameter weights to improve the accuracy when data distribution differs from the source domain to the target domain. Lower (convolutional) layers of CNNs capture more domain-specific information concealed within the image by convolving it with a kernel (or filter), while deeper (dense) layers are responsible for learning information that is relevant for making the decision<sup>37,38</sup>. In the TL framework presented in this work, the source classifier CNN was trained on the CWRU dataset, and the target classifier (TL-CNN) was then allowed to leverage this learned information of decision-making by transferring

Hyperparameter	Range	Optimal value
No. of conv layers	[2, 3]	2
No. of dense layers	[2, 3]	3
Filters (1st Conv layer)	[32, 16, 8]	16
Filters (2nd Conv layer)	[32, 16, 8]	8
Filters (1st Dense layer)	[32, 16, 8]	16
Filters (2nd Dense layer)	[32, 16, 8]	8
Learning rate	[0.001, 0.01, 0.1]	0.01
Batch size	[64, 128, 256]	64
No. of epochs	[10, 20, 30]	30

**Table 1.** Hyperparameter tuning of CNN architecture using *RandomSearch* algorithm.

Layer name	Layer (kernel size × no. of filters)	Output shape	Trainable parameters
L1	Conv (3 × 3 × 16)	128 × 500 × 16	160
L2	Maxpool (2 × 2)	64 × 250 × 16	0
L3	Conv (3 × 3 × 8)	64 × 250 × 8	1160
L4	Maxpool (2 × 2)	32 × 125 × 8	0
L5	Flatten	32,000	0
L6	Dropout (0.2)	32,000	0
L7	Dense (16)	16	512,016
L8	Dense (8)	8	136
L9	Dense (1)	1	9
Total trainable parameters			513,481

**Table 2.** Summary of CNN architecture for scalograms.

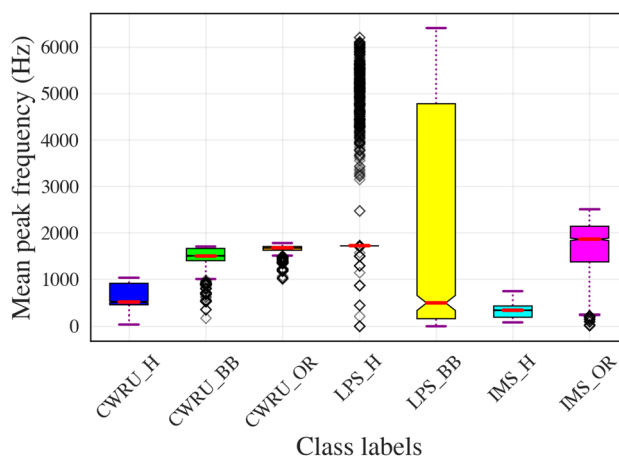
weights of certain layers in the CNN, and retraining the remaining layers with data from the target domain. We explored transferring weights of individual convolutional and dense layers, to study the sensitivity of each layer to the model performance on the target domain. Separately, we also considered TL-CNNs with weights transferred for the dense layers (i.e. all dense layer weights frozen, and convolutional layers re-trained). We considered two different target domain cases for validation of the proposed deep TL architecture i.e., when the target domain dataset is derived from a different machine with low precision sensor (LPS) and when the target domain contains limited and unlabelled data (IMS).

## Results and discussion

Since IMS data contains an outer race (OR) fault and the LPS dataset contains a ball bearing (BB) fault, both these fault conditions were chosen from the CWRU dataset for training CNNs for binary classification of faults. For evaluating the CNNs, we generated 5000 scalograms from each dataset (CWRU, IMS, LPS): 2500 each for healthy and faulty classes. We have used the hold-out cross-validation technique to split the 5000 scalograms from each dataset in the ratio of 50%:20%:30% for training, validation and testing of the CNNs, respectively. Table 3 shows a summary of the results thus obtained. The first four cases in Table 3 i.e., cases (a)–(d), show the model performance when training and test data belong to the same distribution (i.e., same data set). As expected, the accuracy of the CNN models for CWRU and IMS datasets was greater than 99%, owing to the superior performance of CNNs for binary classification of faults in CBM. However, classification accuracy for the LPS dataset i.e., case (d), was lower than CWRU and IMS datasets, which comprise data collected with high-precision sensors. Apart from sensor precision, another factor that contributes to the low classification accuracy in case (d) is the significant overlap in mean peak frequency (MPF) for healthy (H) and ball bearing fault (BB), as seen in Fig. 2. MPF is a key time-frequency domain feature that increases sharply with advancing defects in machines<sup>36</sup>. The MPF for healthy and faulty classes is well-separated in CWRU and IMS datasets. Therefore, it is not surprising that the accuracy and F1 score is lower for case (d) as compared to cases (a)–(c) in Table 3.

Case	Train data	Test data	Trainable parameters	Accuracy (%)	F1 score (%)
Training and testing on same data set					
(a)	CWRU (H, BB)	CWRU (H, BB)	513,481	100	100
(b)	CWRU (H, OR)	CWRU (H, OR)	513,481	100	100
(c)	IMS (H, OR)	IMS (H, OR)	513,481	99.6	99.6
(d)	LPS (H, BB)	LPS (H, BB)	513,481	89.73	88.04
Training and testing on different data sets without TL					
(e)	CWRU (H, OR)	IMS (H, OR)	513,481	50	0
(f)	CWRU (H, BB)	LPS (H, BB)	513,481	50	0
Training and testing on different data sets with TL (TL-CNN)					
(g)	CWRU (H, OR)	IMS (H, OR)	1320	98	98
(h)	CWRU (H, BB)	LPS (H, BB)	1320	96.6	96.6

**Table 3.** CNN performance results on various data sets using scalograms as inputs.



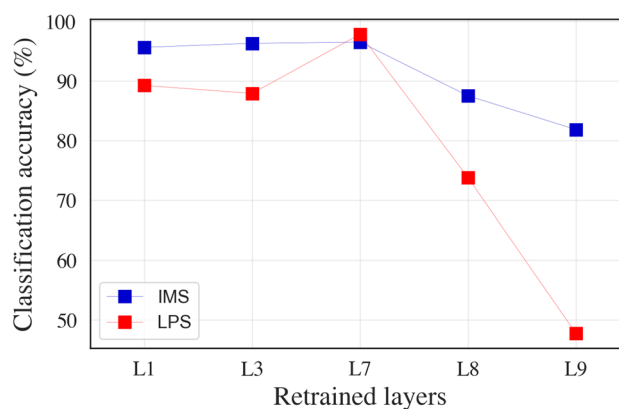
**Figure 2.** Mean peak frequency (MPF) for vibration data in the three data sets (CWRU, IMS, LPS) corresponding to various operating conditions (H healthy, BB ball bearing fault, OR outer race fault).

**Impact of retraining individual TL-CNN layers on accuracy.** For parameter-based TL, the appropriate transfer of weights from the source to the target domain is critical for improving model performance. The lower (convolutional) layers in CNNs capture domain-specific information, and the deeper (dense) layers contribute to the effective learning required for data classification<sup>37,38</sup>. Since the source and target domain datasets vary concerning the environment and load conditions, we hypothesise that the CNN architecture can be adapted to these changes by retraining the weights in only the lower layers, while weights in the deeper layers weights can be transferred unmodified for effective fault classification. To validate this hypothesis, we examined the sensitivity of retraining individual convolutional and dense layers in the CNNs, while keeping the weights of other layers frozen (parameter transfer). The impact of retraining weights of individual layers for IMS and LPS target domain datasets is shown in Fig. 3. We observed that the classification accuracy obtained by retraining lower layers was greater than the conventional method of retraining deeper layers in the model. Along with the improvement in accuracy compared to the conventional approach, the number of retrainable parameters is also significantly smaller, which can certainly contribute to improving the training time for the model.

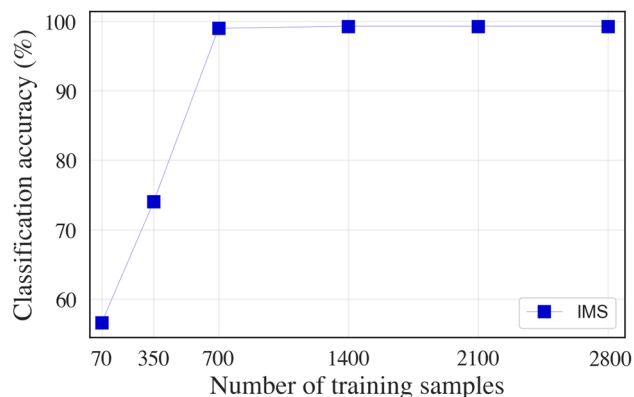
**Performance evaluation of TL-CNN for domain generalisation.** The domain generalisation ability of the CNNs trained on the CWRU dataset was abysmal in absence of transfer learning, resulting in poor accuracy and F1 score: cases (e) and (f) in Table 3. Utilising TL-CNN with weights retrained for convolutional layers L1 and L3 (and frozen for dense layers L7, L8, and L9) resulted in significant improvement in domain adaptability, and high accuracy and F1 score for cases (g) and (h). For instance, the accuracy and F1 score improved from 50% and 0% (i.e. case (e) in Table 3) to 98% and 98% (i.e. case (g) in Table 3), respectively; and from 50% and 0% (i.e. case (f) in Table 3) to 96.6% and 96.6% (i.e. case (h) in Table 3), respectively. The number of trainable parameters for cases (g) and (h) were mere 1320, corresponding to the sum total of the number of trainable parameters in convolutional layers in Table 2.

**Impact of the size of target domain data set.** As previously mentioned, practical installations of vibration CBM systems result in the generation of a disproportionate amount of healthy data as compared to data in faulty operating conditions due to the abrupt and sudden failure of machines, unlike lab-generated datasets with artificially engineered and controllably introduced faults. Thus, the TL-CNNs may have to be retrained with a small number of samples from healthy and faulty operating conditions. To emulate this scenario, we utilised the TL model performance on the run-to-failure IMS dataset that possesses a large number of healthy samples as compared to faulty samples and used MPF as the signature for assigning labels, using the method described by Mukherjee et al.<sup>36</sup>. Figure 4 presents the accuracy of the deep TL-CNN for a varying number of samples used for retraining the convolutional layers. When the model was retrained using 2800 samples (i.e. scalograms), the accuracy obtained was greater than 95%. When half the number of samples were used for retraining (i.e. 1400), the accuracy was still in excess of 95%, thus highlighting the utility of the TL-CNN for target datasets of limited size and significantly different distribution from the source domain dataset. Note that the TL-CNN showed accuracy in excess of 95% even with as low as 700 samples used for retraining, while the performance of the TL-CNN with scalograms degraded significantly below 700 samples.

Table 4 compares various recent parameter-based TL techniques used for domain generalisation for efficient machine fault diagnosis. He et al. have proposed an improved deep transfer autoencoder-based TL method to enhance classification accuracy in diagnosing gearbox faults even in the presence of significant change in the working condition but when the source and target domain data belong to the same dataset<sup>39</sup>. Lu et al. have used an AlexNet-based pre-trained network to automatically identify and classify distinct bearing faults with improved accuracy for varying load conditions utilising the time-frequency features (spectrograms) generated from the raw time-series vibration signals<sup>40</sup>. Yet in another study, a pre-trained VGG19 network has shown adequate prediction to diagnose different types of bearing faults even when the network is trained on the non-manufacturing specific data (ImageNet dataset)<sup>41</sup>. However, as the source domain data used for training the



**Figure 3.** Impact of retraining individual layers in deep TL-CNNs for parameter transfer (while keeping weights of other layers frozen) for scalograms (layer names as per Table 2) used as inputs.



**Figure 4.** Classification accuracy of TL-CNNs (trained on CWRU, followed by re-training and testing on IMS) for the various number of training samples from the target domain used for retraining the convolutional layers.

References	Method	Input features	Dataset	Application scenarios	Performance remarks
He et al. <sup>39</sup>	Improved deep transfer autoencoder	Raw time-series vibration data	Gearbox fault dataset	Same source and target domain datasets but with varying operating condition	90% accuracy when target domain has significant changes in operating conditions
Lu et al. <sup>40</sup>	AlexNet	Spectrograms	Bearing dataset (CWRU)	Source domain: non-manufacturing data (ImageNet), target domain: manufacturing (bearing data)	99.7% accuracy when target domain has significant changes in operating conditions
Wang et al. <sup>41</sup>	VGG19	Scalograms	Bearing dataset (CWRU)	Source domain: non-manufacturing data (ImageNet), target domain: manufacturing (bearing data)	93% accuracy when target domain comprises different fault types and severities
Li et al. <sup>42</sup>	CNN	Raw time-series vibration data	Bearing dataset (CWRU), gearbox fault dataset	Source and target domain datasets with various operating conditions and distinct fault components	> 90% accuracy when all CNN layers are fine-tuned with target domain data (lower accuracy when only last layer is fine-tuned)
Chen et al. <sup>43</sup>	CNN	Raw time-series vibration data	Bearing dataset (CWRU), gearbox fault dataset, and lab generated bearing data	Source and target domain datasets across different machines and varying operating conditions	99% accuracy with all training samples (accuracy number drops with reduced number of training samples)
This work	Light-weight CNN	Scalograms	Bearing dataset (CWRU), run-to-failure bearing data (IMS), lab generated data with low precision sensor (LPS)	Source and target domain datasets across different machines and sensors of different precision	98% accuracy for high precision, unlabelled sensor data in target domain and 96.6% accuracy with labelled, low precision sensor data in target domain, > 95% accuracy even with halving number of samples used for retraining

**Table 4.** Comparison of the proposed work with other parameter-based transfer learning methods used for domain generalisation for efficient machine fault diagnosis.

network are significantly different from the target domain vibration encodings of the machines (spectrograms and scalograms), a large number of trainable parameters are required for domain adaptation thus, increasing computation cost. A joint diagnostic model comprising a CNN wherein the knowledge learned on one fault (bearing) is applied to efficiently predict another fault (gearbox) has been successfully demonstrated by Li et al.<sup>42</sup>. Chen et al. have presented a CNN-based TL model utilising raw vibration data that has shown improved domain generalisation capability when the source and target domain dataset belongs to different machinery and even with varying working conditions<sup>43</sup>. Although a significant improvement in domain generalisation is achieved using the parameter-based TL techniques, in most of the reports, a publicly available vibration dataset acquired using high-precision sensors is used to validate the effectiveness of the models which limits the deployment of most of the models in a real-world scenario of limited access to resource utilisation.

## Conclusion and future work

In summary, in the proposed work we have presented a deep TL-based model for efficient domain generalisation of CNNs trained on a widely accepted benchmark-bearing fault dataset, for high classification accuracy when tested to different distribution target domain vibration CBM datasets obtained from different machines and operating conditions. We studied the performance of the model using a scalogram (time-frequency representation of

vibration signal) as input due to its flexible time-frequency resolution, which offers superior classification results. Besides, the TL-CNNs for scalogram presented in this work utilize only 1320 trainable parameters and are thus suitable candidates for edge-implementation and re-training of these models for autonomous inference at the sensor node e.g. Raspberry Pi single board computer. Additionally, we have demonstrated that deep TL-CNNs can be helpful to improve the accuracy of fault classification for unlabelled datasets with few samples available for training. In future work, we aim to explore methods for data fusion from multiple source domains to enable learning of diverse operating conditions to improve domain generalisation further and mitigate *negative transfer*, a common drawback of TL models, wherein the performance of the model on older data is compromised when retrained on newer data that may have different distribution due to concept drift<sup>44</sup>. The proposed technique of transfer learning allows attaining efficient generalisation capability for distinctly distributed target domain datasets with a lesser number of trainable parameters and even with a lesser number of training samples albeit the training on the source domain should include the fault class of the target domain dataset for effective supervised fault classification. While the time-frequency representations explored in this work yield high classification accuracy with self-learned features, they require computing resources for inference, that may be beyond the reach of low-cost edge devices such as microcontrollers. We shall also explore feature engineering strategies to capture the variance in such diverse data sets for developing edge-based compact models<sup>45</sup>, which could truly make a huge impact on vibration CBM. Along with that, we also aim to study and incorporate explainability into such machine learning models, for providing automated and actionable feedback with limited intervention from domain experts.

### Data availability

The CWRU and IMS datasets analysed in the current study are available at <https://engineering.case.edu/bearingdatacenter> and <https://www.nasa.gov/content/prognostics-center-of-excellence-data-set-repository>, respectively. The LPS dataset analysed in the current study is available from the corresponding author upon reasonable request.

Received: 29 October 2022; Accepted: 20 April 2023

Published online: 24 April 2023

### References

- Bajic, B., Rikalovic, A., Suzic, N. & Piuri, V. Industry 4.0 implementation challenges and opportunities: A managerial perspective. *IEEE Syst. J.* **15**, 546–559 (2020).
- Randall, R. B. *Vibration-Based Condition Monitoring: Industrial, Automotive and Aerospace Applications* (Wiley, 2021).
- Xue, S. & Howard, I. Torsional vibration signal analysis as a diagnostic tool for planetary gear fault detection. *Mech. Syst. Signal Process.* **100**, 706–728 (2018).
- Jan, S. U., Lee, Y.-D., Shin, J. & Koo, I. Sensor fault classification based on support vector machine and statistical time-domain features. *IEEE Access* **5**, 8682–8690 (2017).
- Wen, L., Li, X., Gao, L. & Zhang, Y. A new convolutional neural network-based data-driven fault diagnosis method. *IEEE Trans. Ind. Electron.* **65**, 5990–5998 (2017).
- Kordestani, M., Samadi, M. F. & Saif, M. A new hybrid fault prognosis method for mfs systems based on distributed neural networks and recursive Bayesian algorithm. *IEEE Syst. J.* **14**, 5407–5416 (2020).
- Li, X., Shao, H., Lu, S., Xiang, J. & Cai, B. Highly efficient fault diagnosis of rotating machinery under time-varying speeds using lsimm and small infrared thermal images. *IEEE Trans. Syst. Man Cybern. Syst.* **52**, 7328 (2022).
- Singh, G. & Kazzaz, S. A. S. A. Development of an intelligent diagnostic system for induction machine health monitoring. *IEEE Syst. J.* **2**, 273–288 (2008).
- Janssens, O. *et al.* Convolutional neural network based fault detection for rotating machinery. *J. Sound Vib.* **377**, 331–345 (2016).
- Pan, T., Chen, J., Zhou, Z., Wang, C. & He, S. A novel deep learning network via multiscale inner product with locally connected feature extraction for intelligent fault detection. *IEEE Trans. Ind. Inf.* **15**, 5119–5128 (2019).
- Zhang, S., Zhang, S., Wang, B. & Habetler, T. G. Deep learning algorithms for bearing fault diagnostics—A comprehensive review. *IEEE Access* **8**, 29857–29881 (2020).
- Zhang, W. *et al.* Deephealth: A self-attention based method for instant intelligent predictive maintenance in industrial internet of things. *IEEE Trans. Ind. Inf.* **17**, 5461–5473 (2020).
- Zhao, X. *et al.* Multi-scale deep graph convolutional networks for intelligent fault diagnosis of rotor-bearing system under fluctuating working conditions. *IEEE Trans. Ind. Inform.* **19**, 166 (2022).
- Xiao, D. *et al.* Domain adaptive motor fault diagnosis using deep transfer learning. *IEEE Access* **7**, 80937–80949 (2019).
- Zhao, X. *et al.* Intelligent fault diagnosis of gearbox under variable working conditions with adaptive intraclass and interclass convolutional neural network. *IEEE Trans. Neural Netw. Learn. Syst.* **169**, 104635 (2022).
- Guo, L., Lei, Y., Xing, S., Yan, T. & Li, N. Deep convolutional transfer learning network: A new method for intelligent fault diagnosis of machines with unlabeled data. *IEEE Trans. Ind. Electron.* **66**, 7316–7325 (2018).
- Tong, Z., Li, W., Zhang, B., Jiang, F. & Zhou, G. Bearing fault diagnosis under variable working conditions based on domain adaptation using feature transfer learning. *IEEE Access* **6**, 76187–76197 (2018).
- Qian, W., Li, S., Yi, P. & Zhang, K. A novel transfer learning method for robust fault diagnosis of rotating machines under variable working conditions. *Measurement* **138**, 514–525 (2019).
- 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).
- Shao, S., McAleer, S., Yan, R. & Baldi, P. Highly accurate machine fault diagnosis using deep transfer learning. *IEEE Trans. Ind. Inf.* **15**, 2446–2455 (2018).
- Wen, L., Li, X. & Gao, L. A transfer convolutional neural network for fault diagnosis based on resnet-50. *Neural Comput. Appl.* **32**, 6111 (2020).
- Li, X., Zhang, W., Ding, Q. & Li, X. Diagnosing rotating machines with weakly supervised data using deep transfer learning. *IEEE Trans. Ind. Inf.* **16**, 1688–1697 (2019).
- Shao, J., Huang, Z. & Zhu, J. Transfer learning method based on adversarial domain adaption for bearing fault diagnosis. *IEEE Access* **8**, 119421–119430 (2020).
- Zhang, L. *et al.* A review on deep learning applications in prognostics and health management. *IEEE Access* **7**, 162415–162438 (2019).



25. Looney, M. An introduction to MEMS vibration monitoring. *Analog Dialogue* **48**, 1–3 (2014).
26. Asutkar, S., Korrapati, M., Gupta, D. & Tallur, S. Novel elastomer vibration sensor for machine health-monitoring applications. *IEEE Sens. Lett.* **4**, 1–4 (2020).
27. Zhao, X., Yao, J., Deng, W., Jia, M. & Liu, Z. Normalized conditional variational auto-encoder with adaptive focal loss for imbalanced fault diagnosis of bearing-rotor system. *Mech. Syst. Signal Process.* **170**, 108826 (2022).
28. Zhao, M., Fu, X., Zhang, Y., Meng, L. & Tang, B. Highly imbalanced fault diagnosis of mechanical systems based on wavelet packet distortion and convolutional neural networks. *Adv. Eng. Inform.* **51**, 101535 (2022).
29. Case Western Reserve University Bearing Data Center. <https://engineering.case.edu/bearingdatacenter> (Accessed November 2022).
30. Lee, J. *et al.* Bearing data set. In *IMS, Univ. Cincinnati, NASA Ames Progn. Data Repos. Rexnord Tech. Serv.* (2007).
31. Feng, Z., Liang, M. & Chu, F. Recent advances in time-frequency analysis methods for machinery fault diagnosis: A review with application examples. *Mech. Syst. Signal Process.* **38**, 165–205 (2013).
32. Yuan, L., Lian, D., Kang, X., Chen, Y. & Zhai, K. Rolling bearing fault diagnosis based on convolutional neural network and support vector machine. *IEEE Access* **8**, 137395–137406 (2020).
33. Liu, D., Cheng, W. & Wen, W. An online bearing fault diagnosis technique via improved demodulation spectrum analysis under variable speed conditions. *IEEE Syst. J.* **14**, 2323–2334 (2019).
34. Malviya, V., Mukherjee, I. & Tallur, S. Edge-compatible convolutional autoencoder implemented on FPGA for anomaly detection in vibration condition-based monitoring. *IEEE Sens. Lett.* **6**, 1–4 (2022).
35. Qiu, H., Lee, J., Lin, J. & Yu, G. Wavelet filter-based weak signature detection method and its application on rolling element bearing prognostics. *J. Sound Vib.* **289**, 1066–1090 (2006).
36. Mukherjee, I. & Tallur, S. Light-weight CNN enabled edge-based framework for machine health diagnosis. *IEEE Access* **9**, 84375 (2021).
37. Somu, N., Sriram, A., Kowli, A. & Ramamritham, K. A hybrid deep transfer learning strategy for thermal comfort prediction in buildings. *Build. Environ.* **204**, 108133 (2021).
38. Asutkar, S., Chalke, C., Shivgan, K. & Tallur, S. Tinyml-enabled edge implementation of transfer learning framework for domain generalization in machine fault diagnosis. *Expert Syst. Appl.* **213**, 119016 (2022).
39. He, Z., Shao, H., Zhang, X., Cheng, J. & Yang, Y. Improved deep transfer auto-encoder for fault diagnosis of gearbox under variable working conditions with small training samples. *IEEE Access* **7**, 115368–115377 (2019).
40. Lu, T., Yu, F., Han, B. & Wang, J. A generic intelligent bearing fault diagnosis system using convolutional neural networks with transfer learning. *IEEE Access* **8**, 164807–164814 (2020).
41. Wang, P. & Gao, R. X. Transfer learning for enhanced machine fault diagnosis in manufacturing. *CIRP Ann.* **69**, 413–416 (2020).
42. Li, X., Hu, Y., Li, M. & Zheng, J. Fault diagnostics between different type of components: A transfer learning approach. *Appl. Soft Comput.* **86**, 105950 (2020).
43. Chen, Z., Gryllias, K. & Li, W. Intelligent fault diagnosis for rotary machinery using transferable convolutional neural network. *IEEE Trans. Ind. Inf.* **16**, 339–349 (2019).
44. Shi, G. *et al.* Incremental few-shot semantic segmentation via embedding adaptive-update and hyper-class representation. In *Proc. 30th ACM International Conference on Multimedia* 5547–5556 (2022).
45. Sanchez-Iborra, R. & Skarmeta, A. F. TinyML-enabled frugal smart objects: Challenges and opportunities. *IEEE Circuits Syst. Mag.* **20**, 4–18 (2020).

## Acknowledgements

This work was partially supported by a grant from the Technocraft Center for Applied Artificial Intelligence (TCA2I) at IIT Bombay [Grant DO/2021-TCAI002-008]. S.A. acknowledges the Ministry of Education (formerly Ministry of Human Resource Development), Government of India, for supporting her Ph.D. scholarship. The authors thank Mr. K. P. Ashwin and Mr. Pranjal Tandon at Promethean Energy Pvt. Ltd. for their assistance with collecting the experimental data set for evaluation.

## Author contributions

S.A. and S.T. conceived the experiments, S.A. conducted the experiments. Both authors analysed the results and prepared the manuscript.

## Competing interests

The authors declare no competing interests.

## Additional information

**Correspondence** and requests for materials should be addressed to S.T.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023