UNIVERSIDADE ESTADUAL DE CAMPINAS
Faculdade de Engenharia Mecânica

ISNARDO CADENA RODRIGUEZ

Domain adaptation methodology for gearbox fault diagnosis under variable speed conditions.

Metodologia de adaptação de domínio para diagnóstico de falhas em caixa de engrenagens em condições de velocidade variável.

CAMPINAS
2024

ISNARDO CADENA RODRIGUEZ

Domain adaptation methodology for gearbox fault diagnosis under variable speed conditions.

Metodologia de adaptação de domínio para diagnóstico de falhas em caixa de engrenagens em condições de velocidade variável.

Thesis presented to the School of Mechanical Engineering of the University of Campinas in partial fulfillment of the requirements for the degree of Doctor in Mechanical Engineering, in the area of Solid Mechanics and Mechanical Design.

Tese apresentada à Faculdade de Engenharia Mecânica da Universidade Estadual de Campinas como parte dos requisitos exigidos para a obtenção do título de Doutor em Engenharia Mecânica, na Área de Mecânica dos Sólidos e Projeto Mecânico.

Orientador: Prof. Dr. Milton Dias Junior

ESTE EXEMPLAR CORRESPONDE À VERSÃO FINAL DA TESE DEFENDIDA PELO ALUNO ISNARDO CADENA RODRIGUEZ, E ORIENTADA PELO PROF. DR MILTON DIAS JÚNIOR.

CAMPINAS
2024

Ficha catalográfica
Universidade Estadual de Campinas (UNICAMP)
Biblioteca da Área de Engenharia e Arquitetura
Rose Meire da Silva - CRB 8/5974

C114d

Cadena Rodriguez, Isnardo, 1996-
Domain adaptation methodology for gearbox fault diagnosis under variable speed conditions / Isnardo Cadena Rodriguez. – Campinas, SP : [s.n.], 2024.

Orientador: Milton Dias Junior.
Tese (doutorado) – Universidade Estadual de Campinas (UNICAMP), Faculdade de Engenharia Mecânica.

1. Redes neurais convolucionais. 2. Adaptação de domínio (Inteligência artificial). 3. Maquinas - Monitoração. 4. Falha de sistema (Engenharia). I. Dias Junior, Milton, 1962-. II. Universidade Estadual de Campinas (UNICAMP). Faculdade de Engenharia Mecânica. III. Título.

Informações Complementares

Título em outro idioma: Metodologia de adaptação de domínio para diagnóstico de falhas em caixas de engrenagens em condições de velocidade variável
Palavras-chave em inglês:
Convolutional neural networks
Domain adaptation
Machines - Monitoring
System failures
(Engineering)
Área de concentração: Mecânica dos Sólidos e Projeto Mecânico
Titulação: Doutor em Engenharia Mecânica
Banca examinadora:
Milton Dias Junior [Orientador]
Samuel da Silva
Aldemir Aparecido Cavalini Junior
Robson Pederiva
Gregory Bregion Daniel
Data de defesa: 28-08-2024
Programa de Pós-Graduação: Engenharia Mecânica

Identificação e informações acadêmicas do(a) aluno(a)
- ORCID do autor: https://orcid.org/0000-0001-5706-3318
- Currículo Lattes do autor: http://lattes.cnpq.br/9571402912549268

UNIVERSIDADE ESTADUAL DE CAMPINAS

FACULDADE DE ENGENHARIA MECÂNICA

TESE DE DOUTORADO ACADÊMICO

Domain adaptation methodology for gearbox fault diagnosis under variable speed conditions.

Metodologia de adaptação de domínio para diagnóstico de falhas em caixa de engrenagens em condições de velocidade variável.

Autor: Isnardo Cadena Rodriguez

Orientador: Prof. Dr. Milton Dias Junior

A Banca Examinadora composta pelos membros abaixo aprovou esta Tese:

Prof. Dr. Milton Dias Junior, Presidente
DSI/FEM/UNICAMP

Prof. Dr. Robson Pederiva
DSI/FEM/UNICAMP

Prof. Dr. Gregory Bregion Daniel
DSI/FEM/UNICAMP

Prof. Dr. Samuel da Silva
DEM/UNESP

Prof. Dr. Aldemir Aparecido Cavalini Junior
FEMEC/UFU

A Ata de Defesa com as respectivas assinaturas dos membros encontra-se no SIGA/Sistema de Fluxo de Dissertação/Tese e na Secretaria do Programa da Unidade.

Campinas, 28 de agosto de 2024.

This thesis is dedicated to my parents, Isnardo and Hilda.

ACKNOWLEGEMETNS

I would like to express my gratitude to everyone who directly and indirectly contributed to the conclusion of this thesis. I really want to thank my family, Thiago and Caribe, who always supported me in the difficult moments and helped me to overcome all the uncertainties. My parents, Isnardo Cadena and Hilda Rodriguez, whose work enabled me to accede to university. They unconditionally believe in my potential and encourage me to follow my dreams. My brother Pablo Cadena, who is my best friend and somebody who I trust absolutely.

I want to thank my advisor Prof. Dr. Milton Dias Junior for the technical knowledge and the guidance to conclude this work. I also want to acknowledge the other Mechanical engineering school professors for their teachings.

I am grateful to my friends who gave me laughs and great memories.

I am grateful to the University of Campinas for allowing me to develop myself academically and professionally.

This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Finance Code 001.

RESUMO

As caixas de engrenagens são componentes chave amplamente utilizados para transmissão de movimento e potência. Elas fornecem alta aplicabilidade devido às suas diversas configurações e relações de transmissão. No entanto, ambientes adversos e condições operacionais exigentes, como flutuações de carga e velocidade, podem causar danos aos seus componentes. Portanto, o diagnóstico de falhas é essencial para garantir a disponibilidade e confiabilidade das caixas de engrenagens. Métodos convencionais, como análise de vibração e extração de parâmetros estatísticos, podem apresentar limitações ao diagnosticar sistemas sofisticados sob condições de trabalho variáveis. Métodos de deep learning ganharam popularidade porque podem mapear relações complexas e extrair automaticamente características dos dados. A precisão notável desses métodos foi demonstrada em numerosos trabalhos. Apesar dos resultados promissores das abordagens de deep learning para diagnóstico de falhas em caixas de engrenagens, algumas questões precisam ser resolvidas. Treinar esses modelos requer grandes conjuntos de dados rotulados, que são caros e demorados de obter. Além disso, a maioria desses métodos assume que as distribuições de dados de treinamento e teste são as mesmas. Em cenários industriais reais, variações nas condições de trabalho podem modificar características representativas dos dados, comprometendo diretamente o diagnóstico. Nesse contexto, este trabalho propõe uma nova metodologia de adaptação de domínio para diagnóstico de falhas em caixas de engrenagens sob condições de velocidade variável. A ideia principal é combinar conhecimento prévio sobre assinaturas de falhas e deep learning para melhorar o uso da informação. A assinatura de vibração de falha é obtida calculando os espectros de frequência e envelope no domínio da ordem. Em seguida, algumas bandas são selecionadas com base nas frequências características das falhas. Uma normalização global é aplicada para evitar a perda de características distintivas. Posteriormente, um autoencoder convolucional extrai características discriminativas dos dados normalizados para cada domínio. Análise de correlação é empregada para alinhamento de características e para diagnosticar falhas em diferentes condições de velocidade. A estrutura introduzida foi avaliada com dados experimentais. A metodologia proposta pode ser empregada para clusterização. Os resultados mostram que o método proposto permite a identificação correta de clusters pertencentes à mesma falha em diferentes velocidades, sem a necessidade de rótulos ou do número de classes em cada domínio. Além disso, a adaptação de domínio proposta pode ser empregada para

construir um sistema de diagnóstico baseado em um modelo de reconstrução. Resultados experimentais demonstram alta precisão na previsão de falhas usando dados de diferentes velocidades daqueles empregados para treinamento. Comparações com modelos convencionais de CNN corroboram que o modelo de adaptação de domínio proposto melhora a robustez do diagnóstico.

Palavras-Chave: Diagnóstico de falha em caixa de engrenagens, Adaptação de domínio, Autoencoder convolucional, Análise de correlação, Condições de velocidade variável.

ABSTRACT

Gearboxes are key components widely used for motion and power transmission. They provide high applicability due to their diverse configurations and transmission ratios. However, harsh environments and demanding operating conditions, such as load and speed fluctuations, can cause damage to their components. Therefore, fault diagnosis is essential to ensure the availability and reliability of gearboxes. Conventional methods such as vibration analysis and statistical parameter extraction can present limitations in diagnosing sophisticated systems under variable working conditions. Deep learning methods have gained popularity because they can map complex relationships and automatically extract features from data. The outstanding accuracy of these methods has been demonstrated in numerous works. Despite the promising results of deep learning approaches for gearbox fault diagnosis, some issues need to be resolved. Training these models requires large sets of labeled data, which are costly and time-consuming to obtain. In addition, most of these methods assume that the data distributions of training and testing data are the same. In real industrial scenarios, variations in working conditions may modify representative features of data, directly compromising the diagnostic. In this context, this work proposes a novel domain adaptation methodology for gearbox fault diagnosis under variable speed conditions. The main idea is to combine prior knowledge about fault signatures and deep learning to improve the use of information. The fault vibration signature is obtained by calculating the frequency and envelope spectra in the order domain. Then, some bands are selected based on the characteristic frequencies of the faults. A global normalization is applied to avoid loss of distinctive features. Afterwards, a convolutional autoencoder extracts discriminative features from normalized data for each domain. Correlation analysis is employed for feature alignment and to diagnose faults across different speed conditions. The introduced framework was evaluated with gearbox experimental data. The proposed methodology can be employed for clustering analysis. Results exhibit that the proposed method enables the correct identification of clusters belonging to the same health condition across different speeds without the need for labels or the number of classes in each domain. Furthermore, the proposed domain adaptation can be employed to construct a diagnosis system based on a reconstruction model. Experimental results demonstrate high accuracy in predicting faults using data from different speeds to those employed to train. Comparisons with conventional CNN models corroborate that the proposed domain adaptation model enhances the robustness of the diagnosis.

Keywords: Gearbox fault diagnosis, Domain adaptation, Convolutional autoencoder, Correlation analysis, Variable speed conditions.

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS AND ACRONYMS

ANN – Artificial neural network

BPFI – Ball pass frequency of the inner race

BPFO – Ball pass frequency of the outer race

BSF – Ball spin frequency

CDA – Class distribution alignment

CF – Common factor

CNN – Convolutional neural network

DA – Domain adaptation

DAE – Denoising autoencoder

FEM – Finite element method

FFT – Fast Fourier transform

GMF – Gear mesh frequency

HSIC – Hilbert-Schmidt criterion

MCD – Maximum classifier discrepancy

MDA – Marginal distribution alignment

MMD – Maximum mean discrepancy

PCA – Principal component analysis

PHM – Prognostic and health management

ReLU- Rectified linear unit

RKHS – Reproducing kernel Hilber space

SVM – Support vector machine

VAE – Variational autoencoder

WCSS – Within-clusters sum of squares

CONTENTS

1 INTRODUCTION

The rapid industry development has increased machinery production and applications in every manufacturing field. Rotary machinery is pivotal in several industries, such as material processing, energy production, transportation, among others. Gearboxes are one of the components that are extensively employed for motion and power transmission. Gearboxes are very versatile due to a variety of arrangements of shafts and gears that allows a large range of transmission ratios. This feature makes this type of equipment to be applied in the more diverse scenarios and conditions. In addition, these devices are considered precise and highly functional. However, harsh environments, large speed fluctuations and severe load cycles cause different damages to be developed in the components of the gearbox. These defects lead to the apparatus malfunction, an increase of the vibration level and in some cases to the fault transmission to connected equipment. Therefore, the fault diagnosis is essential for the correct maintenance planning and quick component changing to ensure the availability and reliability of gearboxes. Figure 1.1 shows the most common approaches for gearbox fault diagnosis found in the literature.

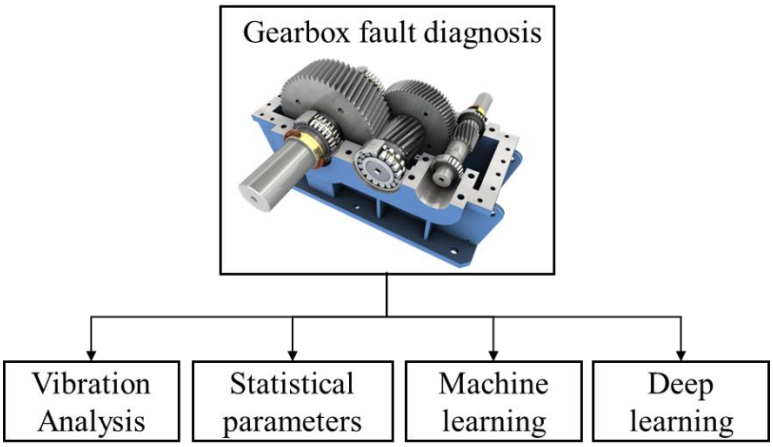


Figure 1.1: Gearbox fault diagnosis approaches.

Most of the fault diagnosis frameworks are vibration based, which means that the gearbox vibration is measured and then studied to infer about the machine health condition. One of the methods that is extensively employed is vibration analysis. In this approach, the collected vibration is treated applying some signal processing techniques to extract relevant information. The Fast Fourier Transform (FFT) is a method par excellence to evaluate the frequency content

of the signal. The spectrum allows to detect peaks in frequencies that are associated with some types of faults. Besides the signal spectrum, other techniques such as wavelet transform, time-synchronous averaging, envelope analysis, signal decomposition among others helps to identify damage indicative patterns. Despite the fact that these analyses can provide valuable insight of possible defects of the gearbox components, some issues need to be considered. Correct fault identification depends on extensive knowledge about the fault as well as clear vibration signature. In addition, factors such as noise and complex gearbox arrangements often affect vibration analysis, making fault recognition difficult and leading to wrong diagnosis. Figure 1.2 presents a frequency spectrum of a machine in healthy state. Peaks at input shaft rotation and two gear mesh frequencies could lead to infer that the device presents any damage. A proper diagnosis then requires a comparative framework to avoid misclassification.

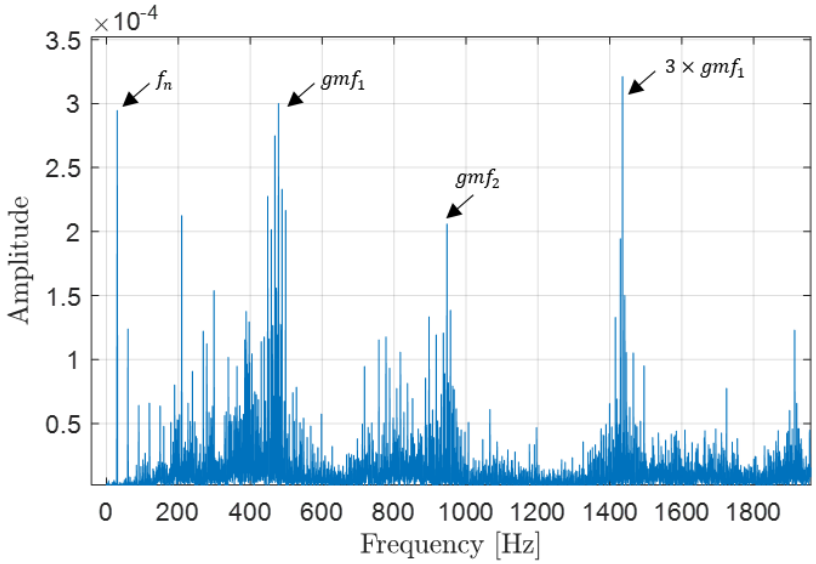


Figure 1.2: Example of a frequency spectrum of a healthy machine.

Statistical analysis is also a widely used tool to investigate anomalies in gearbox components. Parameters such as the root mean square (RMS), shape factor, kurtosis, skewness, peak value, crest factor and others are calculated to assess the measured vibration. Changes in these parameters can indicate changes in the machine health condition. Statistical parameters have proven to be effective to distinguish between a normal state and faulty state in several studies, although when working conditions vary, the diagnosis is compromised. As these parameters are strongly related to the vibration amplitude, they are sensible to outliners and vibration level variation due to load and/or speed fluctuations. Thus, the analysis can provide ambiguous results. On the other hand, statistical features do not provide information about the

fault type and their classificatory power is limited. Figure 1.3 shows nine statistical parameters calculated for six different health conditions. The parameters are as follows: RMS, standard deviation, shape factor, kurtosis, skewness, peak value, impulse factor, crest factor and clearance factor. As illustrated in Figure 1.3, multiple parameters exhibit similar values across various fault patterns. This similarity complicates the determination of whether the observed variations are indicative of a change in the machine's health state or are merely due to noise.

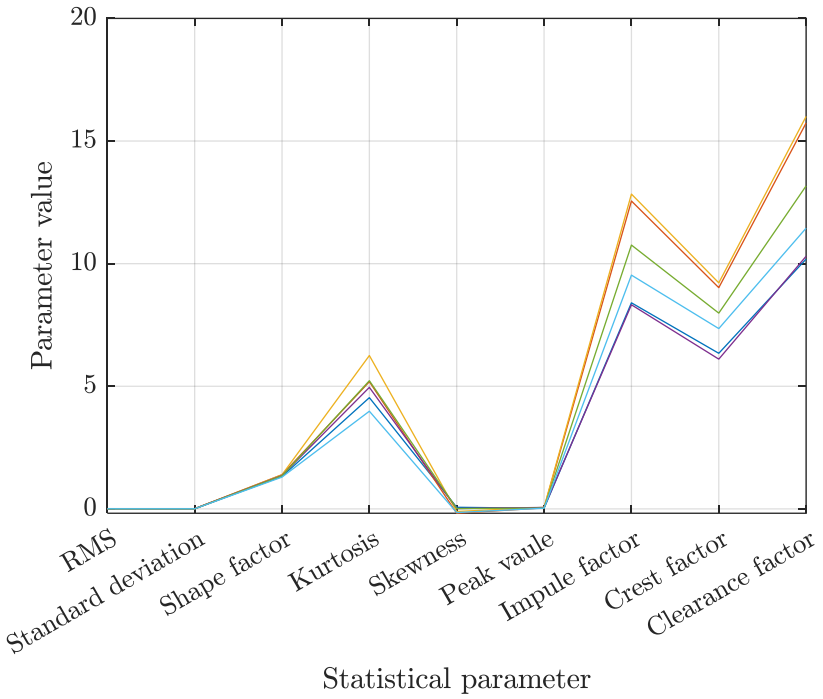


Figure 1.3: Statistical parameters for six different fault patterns.

In recent years, machine learning methods have become protagonists in machinery fault diagnosis. They can be classified into two groups: shallow methods and deep learning methods. In the former group, algorithms such as support vector machine (SVM), random forest and decision tree appear within the most employed approaches. They are useful solving problems where the machine dynamics is not complicated, and environmental conditions are favorable. Nevertheless, dealing with complex scenarios and sophisticated configurations with these methodologies cannot be suitable because of their limited modeling capability. On the other hand, algorithms from the latter group dominated the fault diagnosis field due to their versatility and high performance in handling non-linearities and complicated mapping problems. The artificial neural networks are the most utilized architecture to predict faults in rotary machinery.

Many studies have investigated the capabilities of neural networks and have shown them to be powerful tools for damage detection. However, some issues still need to be solved. Their training demands large sets of labeled data, which are complicated and costly to obtain. In addition, depending on the "depth" of the neural networks, high computational power may be required. Another concern is the generalization of the neural networks. It is well known that their prediction performance with data belonging to the same distribution as the one used in training is outstanding. However, when the data distribution varies, the diagnosis results are highly compromised. Therefore, these problems create the need for new research.

1.1 Motivation

Model generalization has been a constant concern in the development of neural networks. Several works have proposed and discussed methodologies to mitigate the prediction problems caused by overfitting. Figure 1.4 shows a schematic representation of an overfitted and optimal fitted models. The model on the left performs very well on training data, but on testing data it might provide poor results. On the other hand, the model at the right strikes a good balance between accuracy and generalization. One of the most common approaches to enhance generalization is the regularization. Regularization involves techniques that aim to prevent overfitting by adding terms to the loss functions. Popular approaches are the L1 regularization and L2 regularization. These terms in the loss function encourages sparsity in the model parameters and shrinks some coefficients close or equal to zero helping in the model stability. Dropout and batch normalization also are extensively used to improve the prediction performance of the models on test samples. The aforementioned procedures have demonstrated to improve model generalization on samples from the same distribution as the training data. Nevertheless, these methodologies do not perform well when domain changes occur.

In fault diagnosis of rotating machinery, especially gearboxes, speed and load fluctuations can generate differences in data distribution. The vibration amplitude as well as the spectral content vary when working conditions change. These changes produce alterations in the fault vibration signatures, which can affect the diagnosis. Several research in recent years have introduced procedures aiming to overcome these issues. Transfer learning is one approach that helps to reduce the domain differences. Transfer learning applies knowledge learned from a specific task to solve a different but related task. This implies that a pre-trained model can be

fine-tuned for a new case, saving time and resources. For example, a neural network trained using data from a robust mathematical model can be fine-tuned to diagnose faults of real machines using transfer learning strategies. Limitations of this approach include the requirement of labeled data in the new domain to fine-tune the model and the need for training every time the data distribution changes.

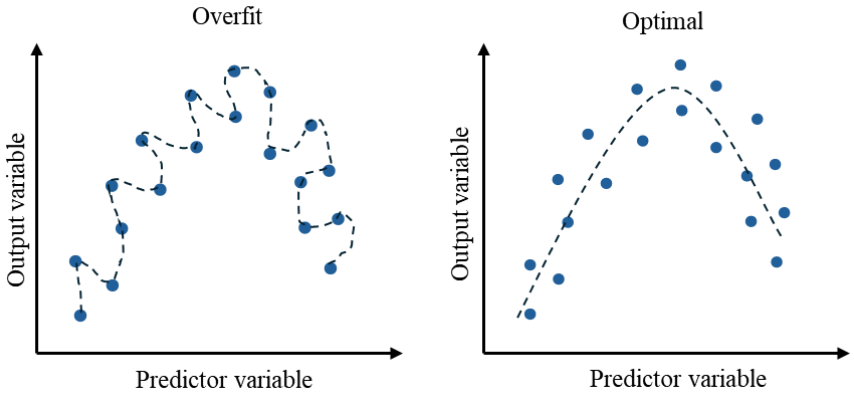
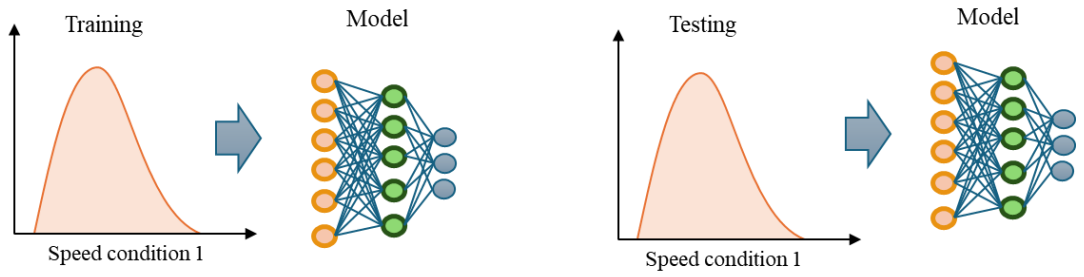


Figure 1.4: Schematic representation of an overfitted and an optimal fitted model.

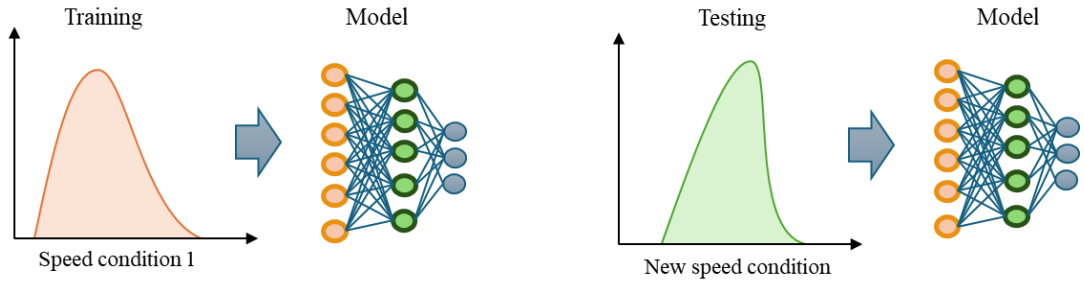
Domain adaptation methods have emerged as a promising solution for model generalization in the presence of domain shifts. Domain adaptation refers to methodologies applied to provide prediction capabilities to deep learning models on data from different distributions. Figure 1.5 depicts an example of a situation where domain adaption can improve the robustness of diagnosis. Traditional deep learning approaches provide high performance on data that present the same distribution that those used for training, for example data measured at the same gearbox speed condition, as shown in Figure 1.5 (a). However, in a more realistic situation, as presented in Figure 1.5 (b) the gearbox speed may vary and it would be desirable for the trained model to provide good diagnosis results in this scenario as well. This is the problem that this work addresses.

Some of the existing domain adaptation methods aim to reduce the discrepancy between the data distribution. For this purpose, the most common approach is to add the maximum mean discrepancy (MMD) term to the model loss function. This results in feature alignment of the source and target data. On the other hand, adversarial training has gained relevance for domain adaptation in fault diagnosis. Feature alignment is achieved by using a domain discriminator whose function is to indicate whether a sample comes from source or target domain. The issue

with the described methods is the limited scope of domain adaptation. They usually are projected to perform well in only one target domain. This implies in new adjustments for every adaptation tasks. Other approach for domain adaptation are the reconstruction methods. They tend to be more adaptable, but their exploitation is still low. In addition, they are usually combined with some statistical metrics to measure the domain discrepancy. This aspect restricts their usage to unsupervised training since the features are modeled as multivariate distributions.



(a)



(b)

Figure 1.5: Deep learning fault diagnosis. (a) Traditional approach using training and testing data from the same distribution. (b) Domain adaptation situation where training and testing data are from different speed conditions.

1.2 Objectives

Considering the context explained before, this work aims to contribute to improving the generalization of machine learning models for gearbox fault diagnosis under variable speed conditions. For this purpose, a novel domain adaptation scheme based on a reconstruction model combined with a correlation analysis is proposed to mitigate the prediction errors in the

presence of speed changes. Prior knowledge about fault vibration signatures is combined with deep learning to optimize the use of the information. The gearbox vibration signal is processed to obtain its frequency content in the order domain. Bands are selected based on the characteristic frequencies of the fault signatures to construct a vector that serves as input to the reconstruction model. A global normalization strategy is applied to preserve fault-distinctive amplitude differences. A convolutional autoencoder is employed to extract domain-invariant features. The correlation analysis is then performed to identify features that represent the same fault type under different gearbox speeds.

1.3 Innovative aspects

This work addresses gearbox fault diagnosis under variable speed conditions using a domain adaptation strategy. Unlike existing methods, where the generalization is achieved increasing the model complexity, the proposed approach combines vibration analysis with deep learning to enhance the use of the information, thereby reducing the required computational power. In addition, no source and target domain labels are required to train the system since this framework uses an unsupervised reconstruction model. The domain adaption is performed by finding correlations between the source and the target domain, avoiding data distribution assumptions. The implementation of the methodology is simplified compared to existing frameworks. Furthermore, the proposed domain adaptation can be applied to a clustering-based diagnosis and to a reconstruction-based fault diagnosis system.

1.4 Organization of the thesis

The remainder of this thesis is organized as follows:

Chapter 2 reviews current works on domain adaptation for gearbox diagnosis, dividing it into three main categories: discrepancy methods, adversarial methods and reconstruction methods.

Chapter 3 provides a theorical background of some of the algorithms used in this thesis, including basic concepts of artificial neural networks, convolutional neural networks, autoencoder and K- means algorithm.

Chapter 4 explains the proposed domain adaptation methodology which is composed of five phases: Fault vibration signature, band selection, normalization, feature extraction and feature synchronization. **Chapter 5** present the data set used to evaluate the proposed domain adaptation framework.

The results and discussions of this thesis are split into two applications for the proposed domain adaptation. **Chapter 6** presents a clustering-based approach for gearbox fault diagnosis under variable speed conditions. **Chapter 7** presents a gearbox fault diagnosis system based on a reconstruction model, and it is compared with traditional convolutional neural network models.

Chapter 8 provides the main conclusions of this thesis and future works suggested by the author.

2 DOMAIN ADAPTATION METHODOLOGIES FOR GEARBOX FAULT DIAGNOSIS: A REVIEW

Domain adaptation (DA) methods mitigate distribution differences between domains, enabling a model trained on source data to perform well on target data. These methods are particularly valuable when there are differences in data distributions between the source and target domains, such as in the case of rotating machinery, where speed and load variation produce changes in fault signatures. Domain adaptation encompasses a range of techniques designed to bridge the gap between domains. These methods can be applied to both shallow and deep approaches, each with its own unique strategies for aligning the source and target distributions. Shallow methods typically involve reweighting instances, transforming features, or sharing parameters to reduce distributional differences. In contrast, deep domain adaptation leverages the powerful representation learning capabilities of deep neural networks to achieve domain invariance, often through adversarial training, reconstruction techniques, or statistical moment matching. In this context, domain adaptation methods for intelligent fault diagnosis can be grouped into three categories: discrepancy-based method, adversarial-based method, and reconstruction-based method.

2.1 Discrepancy-based methods

Discrepancy-based methods evaluate the distance between the source and target domains at the feature layer level of the model and employ statistical techniques to minimize domain differences. The most common approach for comparing and reducing distribution shifts is the maximum mean discrepancy (MMD). MMD evaluates the difference in mean values between the source domain and the target domain. Ghifary et al., (2014) combined MMD with neural networks models for the first time. In fault diagnosis of rotating machinery field, several relevant studies have utilized MMD or its variations in conjunction with deep learning models.

2.1.1 Maximum mean discrepancy metric

The Maximum Mean Discrepancy is defined as the maximum difference between the mean embeddings of two distributions in a reproducing kernel Hilbert space (RKHS). The key idea is to map the samples from the distributions into an RKHS using a kernel function and then compare the means of these mapped samples. In domain adaptation, it is the most widely used metric to compare data distributions.

Lu et al., (2017) introduced a discrepancy term combined with a regularization term to the neural network loss function. The maximum mean discrepancy causes data distributions to become similar by performing domain adaptation, enabling fault diagnosis in different scenarios. The weight regularization term prevents the loss of essential information from the original data during its mapping into the shared subspace, ensuring that different faults remain identifiable after adaptation. Li et al., (2020) employed the maximum mean discrepancy to measure and reduce the disparity of features extracted from the source and target domain in multiple layers of the neural network instead of the final one, focusing on the fully connected ones since the data distribution is retained in these layers, as shown in Figure 2.1. Thus, the domain adaptation task is performed over a larger space producing improved results under domain changes. The method was called Deep Balanced Domain Adaptation Neural Network and required label data from source and target domain. Following the previous idea, Xiong et al., (2021) constructed a multi-block neural network in order to measure the discrepancy of multiple blocks for domain adaptation. Each block was composed of different types and number of layers. In addition, the discrepancy term in the cost function is a sum of central moment discrepancies of each dense block. The methods aimed to enhance feature transferability in deep neural networks for fault diagnosis.

Cao et al., (2020) proposed a deep domain-adaptive multi-task learning model called Y-Net for fault diagnosis in planetary gearboxes, using residual modules to enhance deep feature separability and reduce model redundancy. In addition, a soft joint maximum mean discrepancy was introduced to link two pipelines, shrinking distribution discrepancies of learned features with auxiliary soft labels. Jiao et al., (2020) utilized a one-dimensional residual network for adaptive feature learning, incorporating joint maximum mean discrepancy and adversarial adaptation discriminator to reduce shifts in joint and marginal distributions across different domains, enabling category-discriminative and domain-invariant feature learning for fault

diagnosis. Extensive experiments on planetary gearbox and rolling bearing datasets validated the effectiveness of these approaches. R. Wang et al., (2022) addressed the multi-source domain adaptation. A residual network was employed to learn discriminative features from multiple source and target data. The domain adaptation was performed by applying the multi-kernel maximum mean discrepancy. In addition, a classifier alignment module was trained to improve the prediction accuracy.

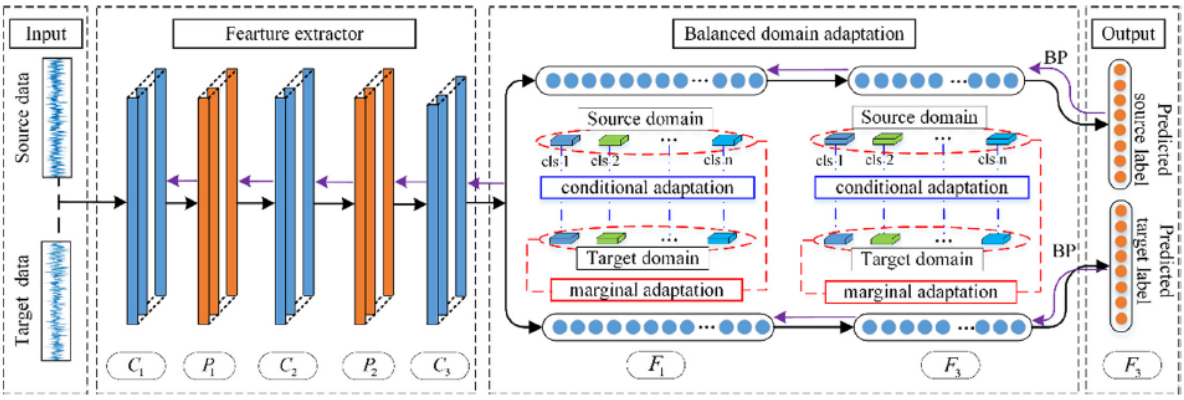


Figure 2.1: Model used by Li et al., (2020) to performed domain adaptation.

The maximum mean discrepancy domain adaptation can also be employed for transfer learning approaches. Scheunemann et al., (2021) combined the Hilbert transform and the fast Fourier transform to obtain the envelope frequency content of raw vibration signals. A convolutional neural network was employed to extract the features. Transfer learning is achieved by adding a layer-wise maximum mean discrepancy term to the loss function in the fine-tuning of the network, that is the discrepancy of the fully connected layers conditioned by predefined bands corresponding to fault locations. C. Chen et al., (2021) used a support vector machine for transfer learning to diagnose gear faults under varying working conditions. Domain adaptation was performed by introducing a large margin projection model, which is a regularization term that measures the maximum mean discrepancy to the SVM.

Qin, Qian, Wang, et al., (2023) investigated the partial domain adaptation to overcome limitations when the source and target domain have different fault types. For this purpose, the global geometrical structure was extracted with an affinity matrix instead of a local geometrical structure obtained by distance weighting. Domain alignment was then performed by maximum mean discrepancy. Manifold regularization enhanced the objective function of partial domain

adaptation by allowing the weight of the geometrical structure to adaptively change based on the feature mapping. Experiments with bearing data demonstrate promising capabilities for partial transfer learning.

J. Li et al., (2024) proposed a network for fault diagnosis of different devices. A multi-domain feature extractor, as shown in Figure 2.2, is constructed by combining different type of neural networks with multi attention to capture key feature information from vibration data in the time, frequency, and time-frequency domains. Thus, the model adaptability to feature from several domain is enhanced. In addition, a bidirectional gated recurrent unit is employed to fuse the extracted features, and the maximum mean discrepancy is used as metric to reduce domain differences and perform transfer learning among different devices.

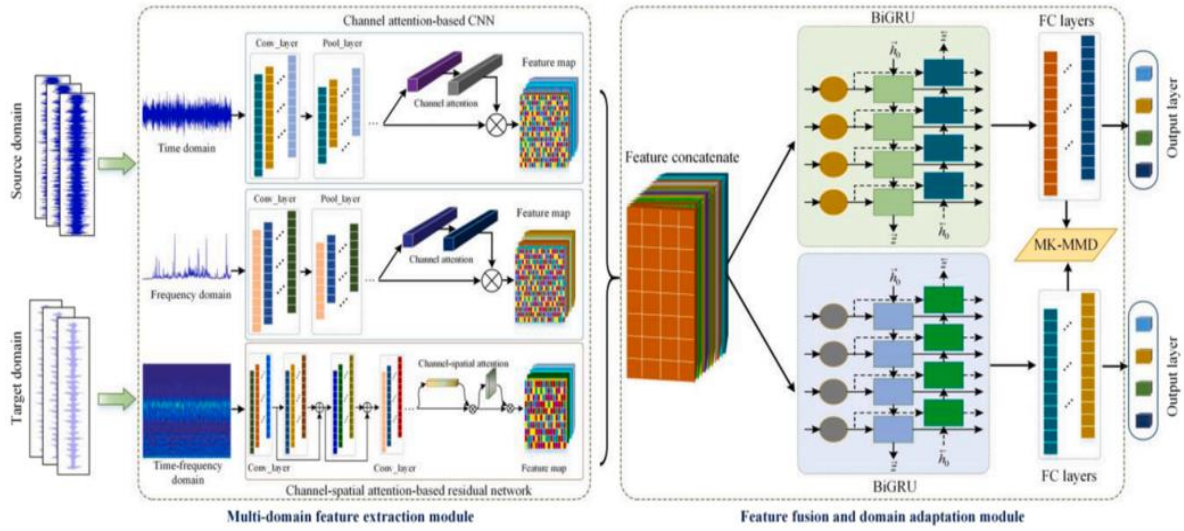


Figure 2.2: Cross-machine fault diagnosis model proposed by J. Li et al., (2024)

2.1.2 Maximum mean discrepancy combined with another metrics

MMD also can be combined with another distance metrics to improve the domain adaption and enhance model generalization, as in the case of Bao et al., (2021) that introduced an enhanced sparse filtering algorithm based on maximum classifier discrepancy (MCD) for transfer fault diagnosis. The Wasserstein distance was leveraged to reduce domain discrepancy between source and target domains. In addition, the algorithm used the probability output

discrepancy of the classifier to identify fuzzy fault samples on the class boundary, enhancing fault diagnosis under speed fluctuation conditions. Rezaeianjouybari & Shang, (2021) reduced the domain difference between multi-source labeled domains and an unlabeled target domain. For this purpose, a sliced Wasserstein discrepancy metric and optimal transport theory are employed to define task-specific bounds. Furthermore, the maximum mean discrepancy appears in the cost function to align the source and target samples. M. Kim et al., (2022) constructed a loss function of a convolutional neural network model adding terms that include maximum mean discrepancy terms, a new semantic clustering term, and a classification lost. The semantic clustering terms bring samples with the same class label closer together and cause samples with different labels to separate. It is based on the pairwise distance metric between each labeled sample from the source domain and is applied at multiple feature levels to obtain robust features with the desired properties. The inclusion of these terms resulted in features from the same class being well-clustered, thereby enhancing the diagnosis.

Jiang et al., (2023) aimed to enhance fault diagnosis performance under gear fluctuating conditions by efficiently extracting hidden features and completing transfer training of data features. For this purposed, the author implemented a combination of maximum mean discrepancy and Wasserstein distance to reduce the sample mean in the regenerative kernel Hilbert space and minimize the average data movement distance. A gradient reversal layer is introduced to improve faults classification and confusion of domain data.

Qin et al., (2023) proposed a domain adaptation mechanism named deep joint distribution alignment, comprising Marginal Distribution Alignment (MDA) and Class Distribution Alignment (CDA) techniques. MDA was enhanced by constructing a statistical metric aligning means and covariances of two domains to reduce marginal discrepancy. CDA utilized Gaussian mixture model and Maximum Likelihood estimation to compute class conditional distributions, matched using information entropy and Wasserstein distance. CDA also mitigates the effects of pseudo labels on the unlabeled target domain, which overcomes some weaknesses of class distribution alignment. A convolutional neural network model was constructed for fault diagnosis based on the marginal distribution and class distribution alignment.

The MMD limitations were analyzed by Qian et al., (2023), who theoretically explored the relationship between Maximum Mean Discrepancy and the kernels. Based on the investigated relation, a maximum mean squared discrepancy metric was introduced. This metric reflects the mean and variance information of data samples in the kernel Hilbert space, enhancing domain confusion. Furthermore, two empirical statistics, biased and unbiased

maximum mean squared discrepancy statistics, were developed to deal with limited sample data in real applications and ensure the effectiveness of maximum mean squared discrepancy. The methos was applied to end-to-end fault diagnosis of planetary gearbox without labeled target-domain samples, showing better discrepancy representation.

2.1.3 Other metrics

Other metrics can be utilized to reduce the domain discrepancies and deal with data distribution changes. Z. H. Liu et al., (2021) constructed an autoencoder network for unsupervised feature learning, enabling the extraction of class-discriminative features from the input data. Then, the discrepancies between the joint learning representations and labels were minimized based on the optimal transport distance. The main idea was to search for a transport plan between the feature and label spaces of the source and target domains while retaining the label information of the source domain. The authors were the first to use the optimal transport theory in domain adaptation for fault diagnosis.

Fan et al., (2022) discussed the development of a Deep Weighted Quantile Domain Adaptation Network to address challenges in fault diagnosis. The network incorporated a weighted quantile discrepancy metric that leverages quantile theory to consider different quantiles' influence on domain adaptation for fault diagnosis. This metric enriched distribution discrepancy measurement methods and enhances distribution matching characteristics with less computational complexity. Su et al., (2023) focused on directly measuring and minimizing decision result matrix discrepancies to facilitate the minimization of distribution discrepancies between two-domain data, enhancing alignment accuracy. The nuclear norm enhanced the precision and robustness of decision discrepancy measurement, mitigating classification errors near decision boundaries.

R. Wang et al., (2024) proposed a novel distance metric addressing conditional distribution discrepancy for cross-domain rotating machinery diagnostics. This metric involved an empirical estimation of the Bures–Wasserstein distance based on the conditional covariance operator, which explicitly constructs the relationship between sample features and labels. Also, geometric features were extracted of different domain data distributions for improved conditional domain alignment. The Bures–Wasserstein distance and the 1-Wasserstein distance were jointly optimized to fully utilize diagnostic knowledge across different dimensions. A

dynamic parameter estimation strategy addressed the hyperparameter selection problem for Gaussian kernel functions. Additionally, the algorithm aimed to minimize transport costs from the source domain sample distribution to the target domain sample distribution, offering interpretability for understanding the cross-domain diagnostic knowledge transfer process.

S. Lee et al., (2024) introduced two main metrics, revolution discrepancy and peak discrepancy, to quantify domain discrepancies between high-speed and very low-speed bearings, considering rotational speed and specification differences. The proposed method consisted of three parts: domain discrepancy quantification, a revolution matching module, and a peak matching module. Revolution discrepancy was utilized in the first training phase to generate revolution-matched target data through the revolution matching model, aiding in extracting speed-invariant features by aligning revolution information. In the second training phase, the peak matching module was implemented to the target data with peak discrepancy to generate peak-matched target data, which is then used for fine-tuning the feature extractor.

2.2 Adversarial-based methods

The core idea behind adversarial domain adaptation is to reduce the discrepancy between the source and target domains by leveraging adversarial training. This involves a game-theoretic approach in which two models, typically a feature extractor and a domain discriminator, are trained simultaneously. The feature extractor aims to generate representations (features) of the input data that are useful for the primary task (e.g., classification) and indistinguishable across domains. The domain discriminator attempts to distinguish whether a given feature representation comes from the source or the target domain. During training, the feature extractor is optimized not only to perform well on the primary task using source domain data but also to trick the domain discriminator into believing that the features it produces are domain-invariant. This is achieved by minimizing the classification loss on the source domain and maximizing the domain discrimination loss, effectively aligning the distributions of the source and target domain features. Some relevant studies that have utilized adversarial learning are explained below.

2.2.1 Adversarial domain adaptation

Jiao, Lin, et al., (2020) presented a double-level adversarial domain adaptation network for cross-domain fault diagnosis. In the diagnostic framework, three players— a feature extractor, a domain discriminator, and a discrepancy discriminator with two label predictors— engaged in two minimax games. In the first game, the domain discriminator tried to distinguish features from the source and target domains, while the feature extractor aimed to fool it. Concurrently, the label predictors were updated to classify source samples accurately. In the second game, the discrepancy discriminator, using the Wasserstein distance, identified ambiguous target samples by maximizing the label predictors' output discrepancy. The extractor, in turn, worked to align target features with the source and deceive the discrepancy discriminator. Zhao et al., (2021) introduced an unsupervised domain adaptation framework called the deep multi-scale adversarial network with attention. The model consisted of two main components: a shared feature generator with multi-scale modules and an attention mechanism, and a fault pattern recognition module with two discriminators. It eliminated the need for labeled information during the transfer process, reducing the time and cost associated with collecting labeled samples in real industry settings. Lou et al., (2022) proposed a fault diagnosis method that utilizes domain adaptation to address the disparity between simulation signals and measured signals. The Finite Element Method (FEM) is used to create a simulation model of the mechanical system, covering various fault types. The FEM simulation signals are adjusted using a generative adversarial network to reduce the distribution gap between FEM and measured fault samples.

Jang & Cho, (2022) proposed a data interpolation method that used a mixing algorithm to fill the high-quality continuous latent space between the source and target domains. An interpolation method was devised employing a reconstruction-based transformation algorithm and a regularization method between the domain data, calculating the regularization factor based on units of reconstruction loss and consistency loss in the trained model. The model presented in Figure 2.3 generated complementary data for both domains, promoting a continuous connection of domain-invariant spaces through interdomain mixing, minimizing domain inconsistency, and generating data that ensured continuous classification performance in the latent space by mixing categories.

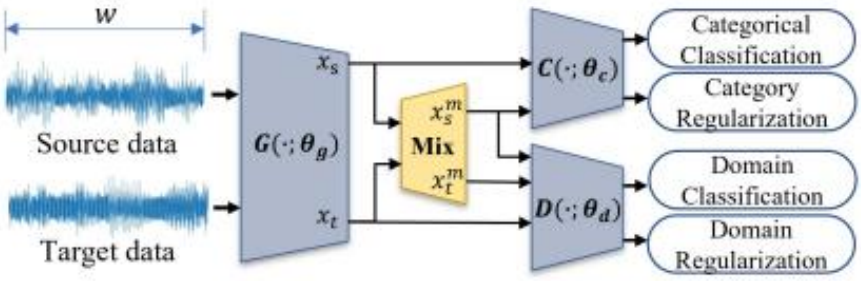


Figure 2.3: Domain adaptation based on a generative model. Source: (JANG; CHO, 2022)

Chen et al., (2023) Introduced an inner adversarial module to extract common features from multi-source domains and enhance domain confusion through dual adversarial training. The authors designed a multi-subnet collaborative decision module for better fusion decisions based on confidence scores computed by the module. This approach addressed the challenge of collaborative fault diagnosis by integrating multi-source domain knowledge effectively. The partial domain adaptation is also considered.

X. Wang et al., (2022) proposed a dual-domain alignment approach that comprised seven modules: two feature extractors, two feature classifiers, a domain discriminator, a source domain alignment module, and a cross-domain alignment module. The source domain alignment maintained the consistency of the source domain features. The cross-domain alignment addressed domain differences that arose by using two fully connected layers to maximize the alignment of the target domain with the source domain.

Tian et al., (2023) proposed an ensemble network framework employing a multi-source mutual-supervised strategy to extract relevant diagnostic knowledge from multiple sources, enabling the identification of both known and unknown health states of target machinery. Furthermore, a transferability metric was introduced. This metric quantified the similarity of each target sample with the known health classes by evaluating and adjusting the consistency, confidence, and entropy of multiple classification results. This process assigned sample-level weights within the adversarial mechanism.

2.2.2 Adversarial training for transfer learning

Adversarial domain adaptation strategies can be combined with transfer learning to improve the model generalization. Z. Chen et al., (2020) utilized two deep encoder networks to create asymmetric mappings to adaptively extract features from raw data in both source and target domains. Additionally, weight transfer and domain adversarial training techniques with inverted label loss were implemented to guide feature learning, aiming to minimize the distribution discrepancy between the source and target domains. Unlike other methods, the authors employed independent encoder network architectures for each domain as shown in Figure 2.4. This design provided flexibility to handle large domain shifts by independently learning domain-specific features.

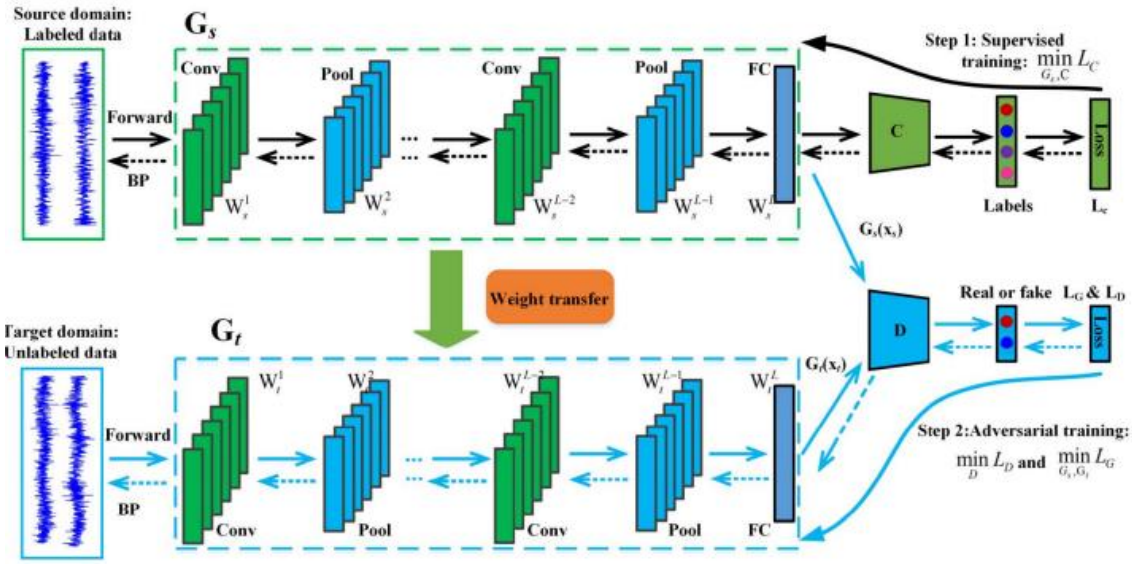


Figure 2.4: Adversarial training model for transfer learning between different domains. Source: (CHEN et al., 2020)

Qin et al., (2021) proposed a parameter sharing adversarial domain adaptation network for fault transfer diagnosis of planetary gearboxes, addressing issues in traditional transfer learning models like high training costs and low classification accuracy. This framework introduced a shared classifier to unify fault and domain classifiers, reducing network complexity. It incorporates the CORAL loss for adversarial training to enhance domain confusion and an unbalanced adversarial training strategy to improve feature extractor domain confusion. Liu et al., (2023) presented a multi-source transfer learning framework called

MADMAN to enhance cross-domain fault diagnosis. MADMAN incorporated three main improvements: the design of the discrepancy matching technique, the construction of the adversarial classifier training method, and the weight adaptive adjustment using the self-attention mechanism. Two classifiers, C1 and C2, were trained to maximize the discrepancy in target features, while the feature extractor minimized the discrepancy to generate consistent features between domains. The discrepancy matching technique effectively reduced the distance of feature distribution between source domains and target domain, as reflected in the MMD values after training with MADMAN being relatively lower compared to other methods. Ran et al., (2024) focused on adversarial domain adaptation for unsupervised scenarios in gearbox fault diagnosis to improve the cross-conditions and cross-machines transfer diagnosis performance. For this purpose, a joint maximum mean discrepancy term was embedded into the loss function to align marginal and conditional distributions. Pseudo-Labels were generated by a self-supervised strategy to facilitate the optimal gradient distribution alignment.

Shi et al., (2022) decomposed the latent spaces into a shared subspace for diagnosis tasks and a private subspace for each domain to capture unique properties, enforcing orthogonal constraints to maintain their independence. Adversarial learning on the shared subspace facilitated the acquisition of generalized diagnostic knowledge from multiple sources, which proved advantageous for the task. This method mitigated the risk of negative transfer at the feature representation level and was anticipated to enhance diagnostic performance on the target task by thoughtfully integrating multiple domains. Furthermore, an instance-entropy-based transfer loss term was introduced into the learning process to minimize the impact of bad samples and bolster condition prediction confidence, thereby further preventing the negative transfer of irrelevant knowledge.

Qian et al., (2024) devised an adaptative intermediate class-wise distribution alignment for mixed transfer tasks. The neural network model utilized a dynamic intermediate alignment layer and introduced Adaptative SoftMax loss with an adaptive decision margin to improve interclass feature separation. The model aligned both global and class-wise distributions of source and target domains, suppressing loss oscillation and slow convergence, and enhancing generalization and robustness. The model only required the SoftMax loss to tune the parameters without any additional discrepancy term. The approach can be applied to domain adaptation and domain generalization without any model modification.

2.2.3 Adversarial training combined with metrics

Some methods experimented the combination of adversarial learning with statistical metrics aiming to extract more specific domain invariant features and minimize the domain divergences. Yu et al., (2021) employed the Wasserstein distance-based asymmetric adversarial domain adaptation for unsupervised domain adaptation in bearing fault diagnosis. This approach integrated a generative adversarial network-based loss and asymmetric mapping to address training difficulties in adversarial transfer learning, especially under serious domain shifts. R. Li et al., (2022) incorporated an asymmetric mapping feature extractor and deep CORAL alignment to extract domain-invariant features and align class-level features between the source and target domains. This approach overcomes the limitations of shared weights in symmetric mapping, allowing the extraction of more specific-domain features. Jia et al., (2022) introduced a distance guided domain-adversarial network that comprises two key modules: a domain-adversarial network and maximum mean discrepancies guided domain adaptation. The authors utilized a stacked autoencoder (SAE) as a feature extractor to learn domain invariant features and MMD to measure non-parametric distance between different metric spaces for improved domain alignment. The method showcased good accuracy and stability under different working conditions.

J. Lee et al., (2024) developed a domain adaptation method for adversarial training with label-aligned sampling. Label-aligned sampling, where the source and target domains were drawn with aligned label distributions, was adopted to mitigate the impact of class imbalance and the domain discriminator shortcut. Using this sampling strategy, samples with the same label distribution were matched in every batch, allowing domain-invariant attributes to be extracted from domain differences rather than from differences in label characteristics, which resulted in improved performance.

2.3 Reconstruction-based methods

Reconstruction models can play a significant role in domain adaptation by learning robust feature representations that help reduce domain differences. The domain difference can be minimized by mapping the source data and the target data, or both domain data, into a shared domain. This can be achieved through various techniques involving encoder-decoder models.

Some approaches involve mapping both source and target domain data into a shared latent space. This shared space enables models to generalize better to the target domain. Variants such as Denoising Autoencoders (DAEs) and Variational Autoencoders (VAEs) can help in learning robust representations that are less sensitive to domain-specific noise or variations, thereby reducing domain discrepancy. Some relevant works that leverage the advantage of reconstruction methods are detailed below.

He et al., (2019) introduced an improved deep transfer auto-encoder depicted in Figure 2.5 for fault diagnosis of different gearboxes under variable working conditions with small training samples. The multi-wavelet activation function was employed for effectively learning features hidden in non-stationary vibration data, enhancing the analysis performance. The correntropy was utilized to modify the cost function, improving the reconstruction quality of the auto-encoder. The pre-trained autoencoder with data from the source domain was fine-tuned using small samples from the target domain. He et al., (2020) improved the previous work by selecting high-quality samples based on a similarity measure using coefficients calculated from the spectrum aiming to pre-train the autoencoder with data which share similar features with the target domain.

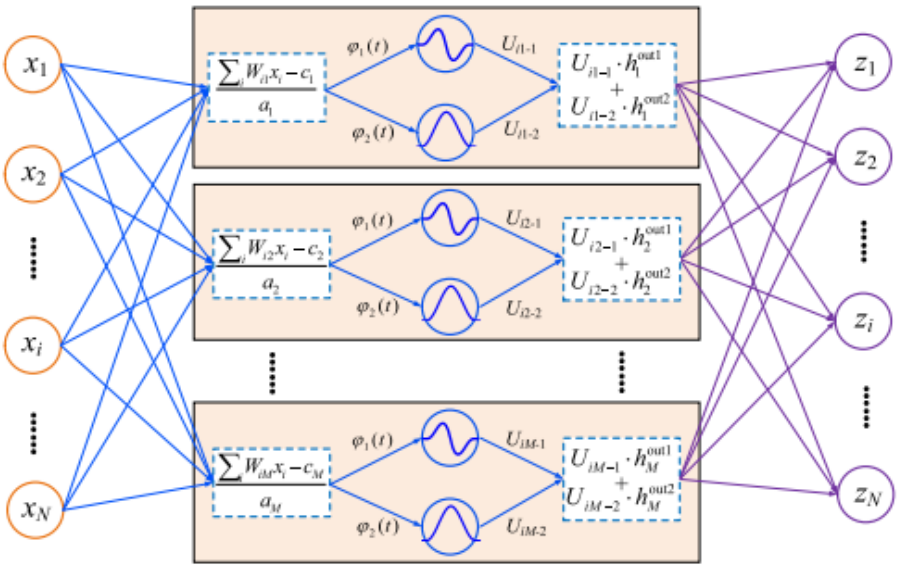


Figure 2.5: Wavelet-based autoencoder for cross-machine gearbox fault diagnosis. Source: (HE et al., 2019), (HE et al., 2020)

X. Li et al., (2020) employed a deep learning-based method that projects different equipment features into the same subspace using an auto-encoder structure, setting the maximum mean discrepancy to measure data distribution discrepancies between different machines. Xiao et al., (2021) proposed a noisy domain adaptive marginal stacking denoising auto-encoder for domain adaptation between different noise levels, integrating Transfer Component Analysis for dimensionality reduction using acoustic signals. The method minimized the maximum mean discrepancy distance between different domains using a forward closed-form solution instead of traditional back propagation algorithm, yielding significant improvements in accuracy and efficiency.

Yang et al., (2022) utilized a Dynamic Domain Adaptation method to address domain shift and mismatch due to different working conditions, consisting of multi-kernel marginal distribution adaptation and multi-kernel conditional distribution adaptation by calculating a dynamic domain factor based on the maximum mean discrepancy. The network incorporated multiple attention mechanism modules to enhance the richness of deep features and promote adaptability in different diagnosis tasks.

Autoencoders architectures can be improved by adding some metrics to the cost functions aiming to extract consistent features across domain variations. Zhang et al., (2020) applied domain adaptation to the sparse filtering algorithm. For this purpose, the authors utilized the L1-norm and L2-norm for the maximum mean discrepancy. Besides, the final objective function obtained by integrating specific equations. Kim & Lee, (2023) employed an expectation-maximization adversarial autoencoder for feature extraction and subspace mapping. A gaussian mixture model was used to cluster the extracted features obtaining multi-variate distributions for each class. The relationship between the source and target domain clusters was inferred utilizing the symmetric Kullback-Leibler divergence metric, where clusters with the smallest probabilistic distance were assigned to belong to the same health condition.

K. Zhao et al., (2023) presented an indirect alignment idea to align the feature distributions of source and target domains in the latent feature space using a Gaussian prior distribution. The Wasserstein Distance was employed as an optimal transport problem to find an optimal transport strategy. It helps transform the source data distribution into an approximate normal distribution, maintaining the geometric characteristics of the distribution during the transformation process. This approach enhances feature alignment by indirectly aligning the distributions, improving the transfer of knowledge from multiple source domains to the target

domain. Pang, (2024) minimized the dependence between extracted features and domain labels using Hilbert-Schmidt Independence Criterion (HSIC) by stacking multiple Maximum Independence Autoencoders (MI-AEs) and fine-tuning them with labeled source samples and normal pattern samples from supporting domains. HSIC was used as a measure of dependence between two data sets to learn domain-independent features in deep transfer models, which is a novel application of HSIC in this context. This approach achieved domain generalization to various working conditions, designing an effective way to learn domain-adapted features.

Variational autoencoders are also employed for cross domain fault diagnosis. Wen et al., (2023) proposed a deep clustering network called clustering graph convolutional network with multiple adversarial learning for fault diagnosis of various bearings. The methodology involved the use of multiple representations of datasets extracted by an autoencoder and graph convolutional network to recognize different classes and their related information. Yuan et al., (2024) utilized a variational auto-encoder (VAE)-based multisource deep domain adaptation model using optimal transport for cross-machine fault diagnosis of rotating machinery. The optimal transport distance was employed to reduce domain drifts across machines, and a multicategory label discriminator for label prediction. This methodology addressed the cross-machine fault diagnosis problem by leveraging optimal transport theory to reduce domain drifts, showcasing accuracy in fault diagnosis.

3 THEORICAL BACKGROUND

Artificial neural networks (ANN) are computational models inspired by the human brain synapsis. They are powerful at pattern recognition, data analysis and modelling complex systems. In recent years, neural network models have gained importance in fault diagnosis due to their versatility, outstanding potential for real applications and high performance compared with traditional methods. Convolutional neural networks (CNNs) stand out among the deep learning models for their ability to automatically extract features, improving the robustness of diagnosis. In addition, kernel utilization enables small input variation without affecting the prediction. This work leverages the advantages of the convolutional neural networks and employs a convolutional autoencoder for the gearbox fault diagnosis. This chapter provides a brief description of the algorithms employed in this thesis including basic concepts of neural network architectures such as the artificial neuron, activation functions and their training. Besides, the convolutional neural networks and convolutional autoencoders are explained. At the end of the chapter, a concise explanation of the K-means clustering algorithm is provided.

3.1 Artificial neural networks: Basic concepts

3.1.1 Artificial neuron

The basic constructive unit of artificial neural networks is the neuron. It was first introduced by McCulloch & Pitts, (1943) for binary entries and later generalized for real values by Rosenblatt, (1958) in the paper called “perceptron’. The perceptron is considered the first artificial neural network and is still widely used in different artificial neural network applications (DA SILVA et al., 2016). Figure 3.1 presents a representation of an artificial neuron.

The set $\{x_1, x_2, \dots, x_m\}$ denotes the input signals and $\{w_1, w_2, \dots, w_m\}$ are the synaptic weights of the synaptic junctions. The relevance of each $\{x_i\}$ is determined by multiplying them by the corresponding $\{w_i\}$, and the activation potential u is a weighted sum of the inputs and

constant b called bias. Thus, the activation potential is sent to an activation function $g(\cdot)$ generating the output y. The bias increases or decreases the degree of freedom of the activation function and allows a non-zero output when inputs are null.

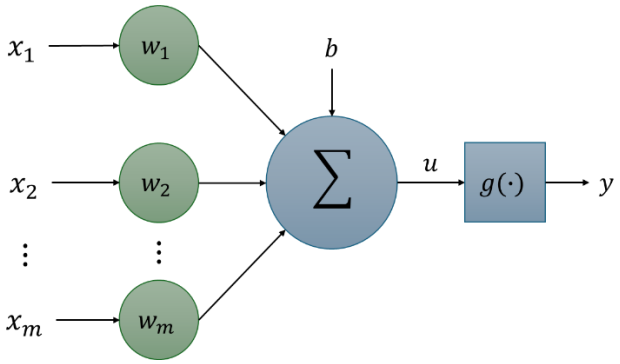


Figure 3.1: Schematic representation of an artificial neuron.

The activation function limits the output values within a reasonable range defined by its functional mapping. Eq. (3.1) and (3.2) synthesize the result produced by an artificial neuron (DA SILVA et al., 2016).

$$u = \sum_{i=1}^{m} w_i \cdot x_i + b \tag{3.1}$$

$$y = g(u) \tag{3.2}$$

3.1.2 Activation functions

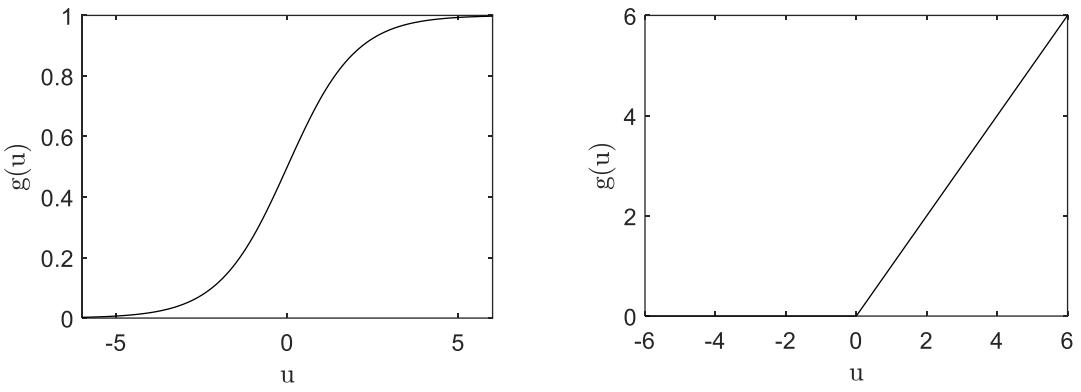
Activation functions influence the flexibility and capability of the ANNs in modelling complex outputs. Besides, activation functions introduce non-linearities to the neural networks avoiding falling into linear regression models. Figure 3.2 shows two activation functions employed in this work.

Sigmoid function also called sigmoidal curve or logistic function is presented in eq. (3.3). The output result produced by sigmoid function will always be real values within the interval $[0, 1]$.

$$g(u) = \frac{1}{1 + e^{-u}} \tag{3.3}$$

The rectified linear unit activation (ReLU) function has become very popular in recent years. This activation function yields a linear mapping for input values greater than zero, while returning zero for negative values and can be expressed as follow:

$$g(u) = \max(0, u) \tag{3.4}$$



(a) (b)

Figure 3.2: Activation functions. (a) Sigmoid function. (b) ReLU function.

3.1.3 Neural network structure

Artificial neural networks are built by connecting several neurons arranged in structures called layers. The layers direct the synaptic connections of the neurons and can be classified regarding their localization in the neural network structure. The input layer is responsible for receiving information coming from the external environment. Data is usually normalized within a defined interval aiming to produce stable mathematical operations performed by the network.

The hidden layers perform most of the data processing and are responsible for extracting patterns associated with the analyzed system. The output layer is responsible for producing the outcome derived from the processing of previous neurons. Figure 3.3 depicts a neural network with one hidden layer.

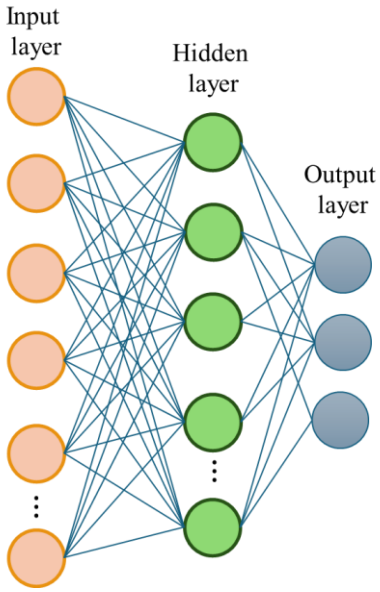


Figure 3.3: Basic structure of an artificial neural network.

Feed forward neural networks with multiple layers are one of the most common architectures employed in machine learning approaches. Information is propagated in only one direction from the input layer to the output layer. These networks are composed of one or more hidden layers. The number of neurons on the input layer must be equal to the length of input data. on the other hand, the number of neurons in the last layer depends on the desired outputs for the analyzed problem. The number of hidden layers and their respective number of neurons is tied to the nature and complexity of the problem as well as the availability of data to model the system. The more hidden layers, the deeper the network is, and the more complex problems can be mapped.

Another type of neural network architecture widely used is recurrent or feedback networks. The main difference to the previous networks is that the output of some neurons can be feedback to the inputs of neurons in certain layers. This type of network is not addressed in this work.

3.1.4 Training the neural networks

The main objective of the training process is to optimize the biases and the connection weights between the neurons. To achieve this, an objective function is defined. The objective function quantifies the error generated by the ANNs through the training. Common objective functions found in the literature include the mean squared error (MSE) defined in eq. (3.5), the mean absolute error (MAE) detailed in eq. (3.6) and the cross entropy shown in eq. (3.7)

$$MSE = \frac{1}{n_s}\sum_{i=1}^{n_s}(y_i - \hat{y}_i)^2 \tag{3.5}$$

$$MAE = \frac{1}{n_s}\sum_{i=1}^{n_s}|y_i - \hat{y}_i| \tag{3.6}$$

$$L = -\frac{1}{n_s}\sum_{i=1}^{n_s}[y_i \log(\hat{y}_i) + (1 - y_i)\log(1 - \hat{y}_i)] \tag{3.7}$$

where n_s is the number of training examples, y_i is the real value and $\hat{y}_i$ is the predicted value. These functions serve as benchmarks for assessing the network's performance and guiding the training towards minimizing error.

Neural networks are trained using the backpropagation algorithm which is a gradient estimation method. Here is a concise overview: for each input-output pair $\{x_d, y_d\}$, where x_d represents a training data example and y_d is its corresponding true label, the network calculates the forward phase, it means, the estimation error given by the loss function L. The next step is the backward phase, where the influence of each network parameter in the produced error is assessed. This is done by computing the derivates of the loss function with respect to each weight w_{jk}^p and bias b_k^p, where p is the number of the layer, k is the number of the neuron in layer p, and j is the number of neural connections. The error is then backpropagated from the output layer to the hidden layers using the chain rule. Since the input layer lacks tunable parameters, the error propagation is limited to the first hidden layer. Finally, each parameter is updated as follows:

$$\left(w_{jk}^{p}\right)^{n+1} = \left(w_{jk}^{p}\right)^{n} - \eta \frac{\partial L}{\partial w_{jk}^{p}} \tag{3.8}$$

$$\left(b_{k}^{p}\right)^{n+1} = \left(b_{k}^{p}\right)^{n} - \eta \frac{\partial L}{\partial b_{k}^{p}} \tag{3.9}$$

where η is the learning rate, the negative signal modifies the searching direction to minimize the objective function and n is the iteration number. This iterative process helps refine the network's performance over subsequent training epochs. Figure 3.4 presents an example of the value of the loss function as function of the number of epochs for the training of a neural network.

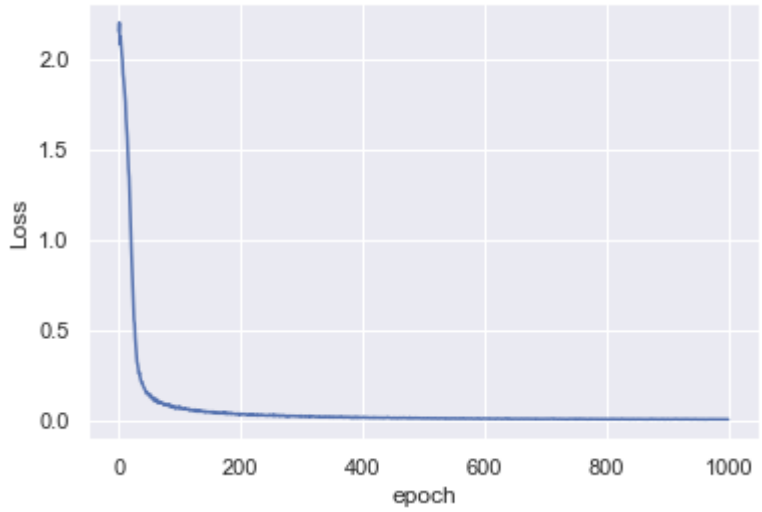


Figure 3.4: Loss function values across several training epochs.

3.2 Convolutional neural networks

A convolutional neural network is a type of deep learning model that extracts information from visual data (LECUN et al., 1998). Deep convolutional neural networks are designed to perform two main tasks: feature extraction and classification.

Feature extraction refers to the process of transforming raw data into numerical representations, maintaining the essential information embedded in the original dataset. Feature

extraction helps in dimensionality reduction, then the data is processed more efficiently. On the other hand, classification involves assigning labels to the data using the extracted features in the previous phase. Figure 3.5 presents a common architecture of a convolutional neural network composed of three types of arrangements: convolutional layers, pooling layers, and fully connected layers.

Hyperparameters of the convolutional neural networks are usually chosen by the user depending on the specific application. The convolutional neural networks are versatile since the changes on the hyperparameters produce a new model providing different results. The first hyperparameter is the number of layers that is commonly called the depth of the network. This is a critical parameter since a deeper network can learn more complex features and patterns from the data, but it is also more susceptible to overfitting. Thus, finding a balance between the number of layers and the complexity of the problem is crucial. Another aspect to decide is the type of layers employed to construct the network. Depending on the arrangement of layers, some networks can be more efficient to specific tasks in comparison to others.

The kernel of filter size is another hyperparameter. A filter is a matrix of weights used to convolve with the input. During convolution, the filter measures how closely a patch of the input matches a particular feature, such as a vertical edge, an arch, or any other shape. Using a larger filter size allows the network to capture more information from the input data, but it also increases the number of parameters. On the other hand, a smaller filter size can reduce the number of parameters, but it might miss some important features in the data.

The stride is a hyperparameter that dictates how many positions the filter shifts across the input data. Utilizing a larger stride can decrease the size of the resulting feature maps, although it may also cause some information to be lost. On the other hand, a smaller stride retains more information but requires more computation time and memory. Therefore, a suitable stride balances between minimizing information loss and maintaining computational efficiency.

Padding is a method used to maintain the spatial dimensions of an input data when applying convolutional layers. This process involves adding zeros around the edges of the input data to create a padded version, which can then be convolved with the filter. Padding helps retain edge information and prevents the loss of spatial resolution. However, it also results in higher memory usage and increased computation time for the network.

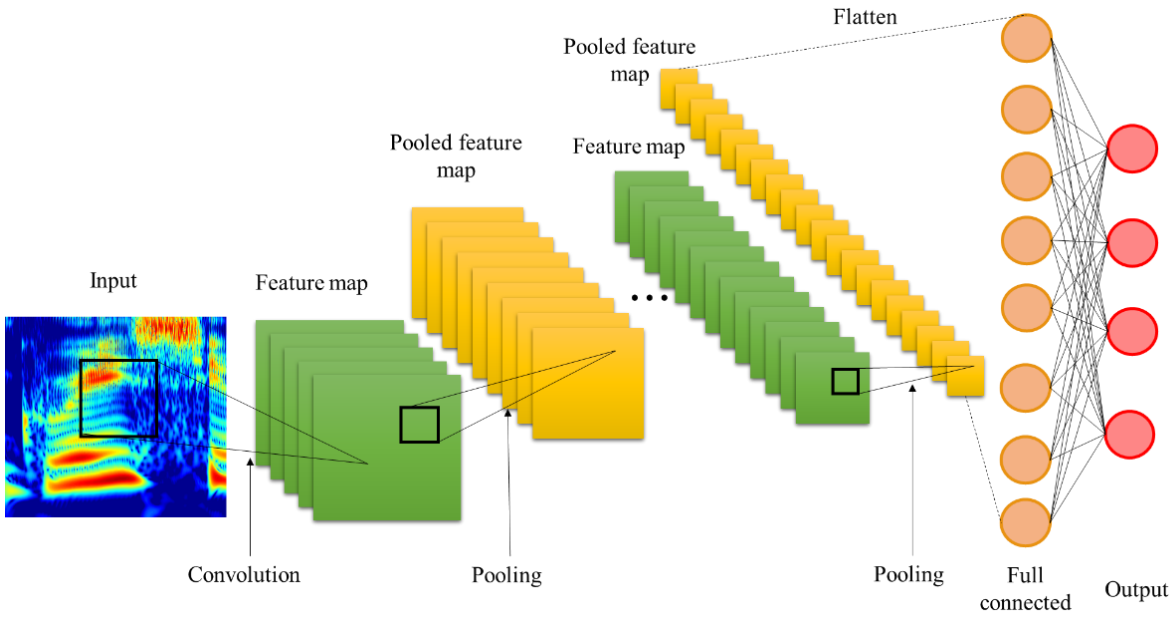


Figure 3.5: Example of structure of a convolutional neural network.

3.2.1 Convolutional layer

Convolutional layers are the core building blocks of CNNs. It comprises a set of filters or kernels, the parameters of which are refined through the training. Convolutional layers take the inputs and produce feature maps to be fed into subsequent layers. Feature maps are generated by kernels through the convolution operation.

Consider a input of size $I = [I_x][I_y][I_z]$ and a kernel of shape $K = [K_x][K_y]$, the feature map is calculated with the following equation.

$$O[z][x][y] = B[z] + W \tag{3.10}$$

$$W = \sum_{c=0}^{I_z} \sum_{k_x=0}^{K_x-1} \sum_{k_y=0}^{K_y-1} I[c][k_x + S_x y][k_y + S_y y] \times K[z][c][k_x][k_y] \tag{3.11}$$

where B is the bias, z the number of channels of the feature map and S_x, S_y are the strides in each direction. The feature map size can be calculated as

$$O_{\{x,y\}} = ((I_{\{x,y\}} - K_{\{x,y\}} + 2P)/S_{\{x,y\}}) + 1 \tag{3.12}$$

Strides dictate how the filters move in each direction over the input. The larger the strides, the smaller the output shape. The number of filters or kernels dictates the depth O_z of the feature map. P denotes padding the output features map with zeros, ensuring that input and output dimensions are equal. Figure 3.6 presents an example of a convolution of a one-channel input using a 3×3 kernel.

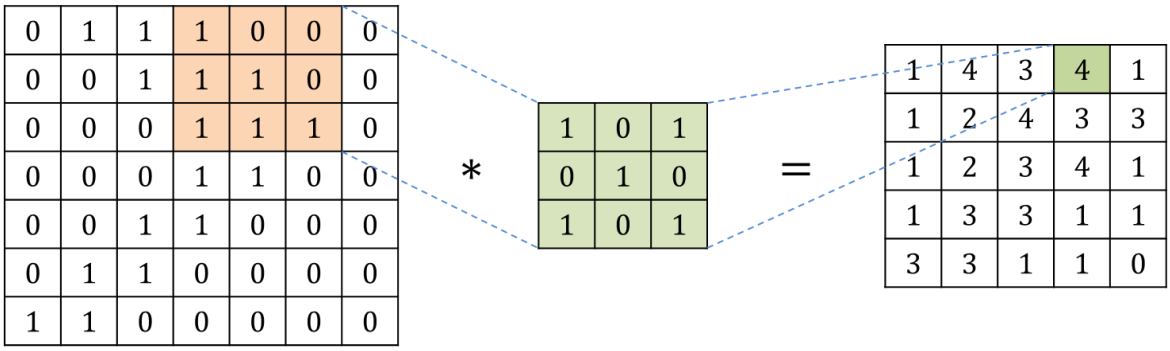


Figure 3.6: Convolution of one-channel input with a 3×3 kernel.

The feature map generated by the convolution is fed into an activation function. The rectified linear unit described in eq. (3.4) is usually employed in convolutional layers. The previous mathematical expressions for convolutional layers can be extrapolated to other input sizes.

3.2.2 Pooling layer

The pooling layer aggregates the output from the previous layer through a pooling window (K_{px}, K_{py}). The max pooling operator is usually used in CNNs. It computes the

maximum over this window and down samples the output using the max value. Figure 3.7 provides an example of a max pooling operation. On the other hand, the average operator computes the average value over the window and down sample the output using this value. As in convolutional layers, the pooling operation can be performed using strides. The main objective of the pooling layer is to reduce the dimension of the feature maps and therefore reduce the number of parameters to train. It is important to highlight that the pooling operation is applied separately to each feature map. This implies that if the preceding feature map has depth z, the pooling layer will generate a feature map with the same depth z.

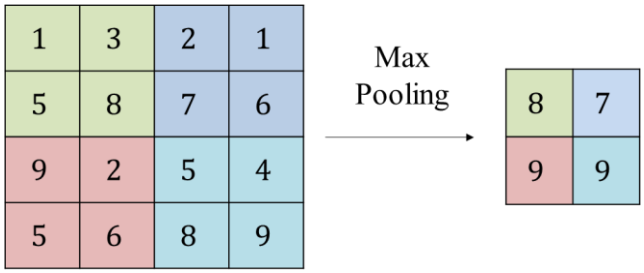


Figure 3.7: Example of Max-pooling operation.

3.2.3 Fully connected layers

Following a series of stacked convolutional and pooling layers, the resulting output feature map is flattened into a vector. This vector serves as the input for a fully connected layer. The data processing follows the explanation given in section 3.1.1. Finally, for multiclass classification, the label of each input data can be attributed applying the following expressions:

$$\hat{y}_i = softmax\left(y_i^{FC}\right) = \exp\left(y_i^{FC}\right) / \sum_{c=1}^{C} \exp\left(y_c^{FC}\right) \tag{3.13}$$

$$label = \arg\max(\hat{y_c}) \tag{3.14}$$

where y_i^{FC} is the output value of i-th neuron in the last fully connected layer, and C the number of classes.

3.3 Convolutional autoencoders

Autoencoders are a type of neural networks structures that efficiently learn representations of input data with no need of labels. The main objective of autoencoders is dimensionality reduction through unsupervised learning. Autoencoders consist of three main parts: the encoder, the bottleneck and the decoder. The encoder is responsible for gradually capture the relevant features of the input data and reducing its dimension until obtaining a compact representation. The final layer of the encoder is called the bottleneck, where the number of neurons is small enough to represent the compressed encoded data without losing significant information. The decoder takes the low dimensional data from the bottleneck and gradually expands them back to their original size. The output layer of the decoder produces a reconstructed output which ideally mirrors the input data. Figure 3.8 presents an schematic representation of an autoencoder.

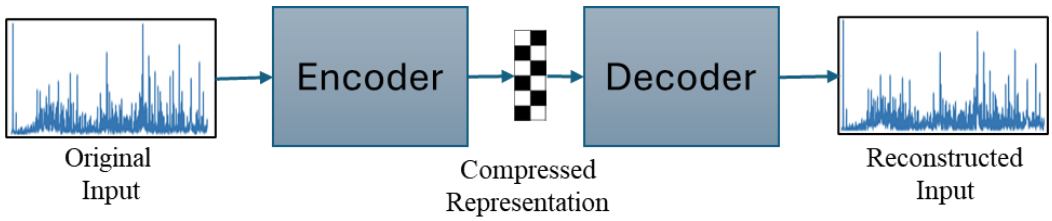


Figure 3.8: Basic structure of an autoencoder.

Convolutional autoencoders are a special type of convolutional neural network extensively used for feature extraction. Convolutional autoencoders are composed of a combination of convolutional and pooling layers in the encoder. In this work, the decoder is composed of transpose convolution and Up sampling layers.

The transposed convolutional layer performs the same operation as a standard convolutional layer but executed in reverse direction. Rather than sliding the kernel over the input and performing element-wise multiplication and summation, a transposed convolutional layer slides the input over the kernel, carrying out element-wise multiplication and summation. This process produces an output larger than the input, with the output size determined by the

layer's stride and padding parameters. Figure 3.9 presents an example of a transpose convolution using a 2×2 filter with stride 1 and padding 0.

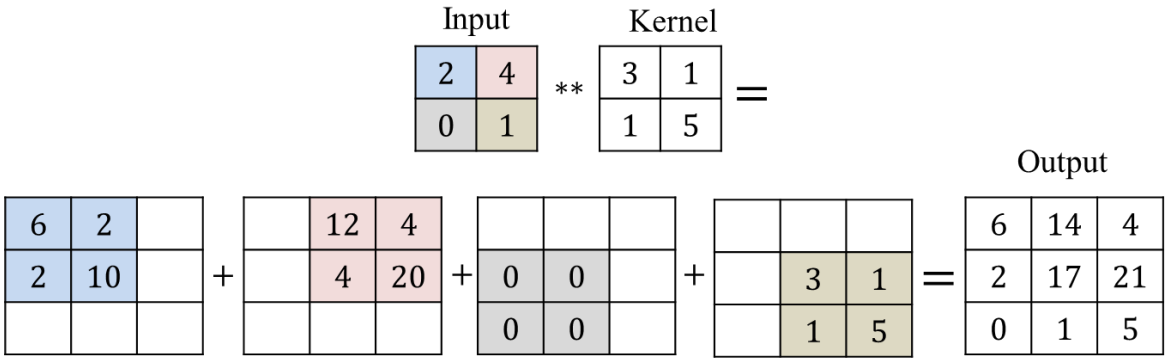


Figure 3.9: Transpose convolution.

Up sampling layers are also employed to increase the dimension of the feature maps but without any trainable parameter. These layers can be thought of as the contrary of pooling layers. Figure 3.10 depicts an example of a up sampling operation.

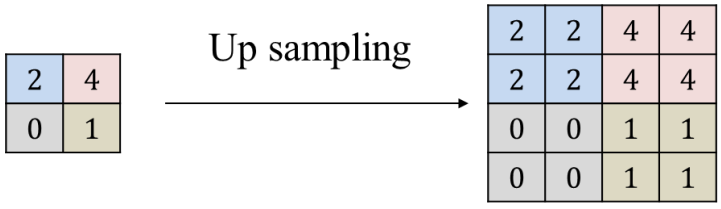


Figure 3.10: Up sampling operation.

The bottleneck can be a convolutional or a fully connected layer. Figure 3.11 presents an architecture of a one-dimensional autoencoder where the bottleneck is a fully connected layer. The most common activation function used in a convolutional autoencoder is the ReLU. However, if the bottleneck is fully connected, the sigmoid function provides better performance. A usual cost function employed in the training of convolutional autoencoders is the mean squared error or some variation thereof. In this work, a one-dimensional autoencoder is utilized to process the gearbox fault signature to extract relevant features with classification potential.

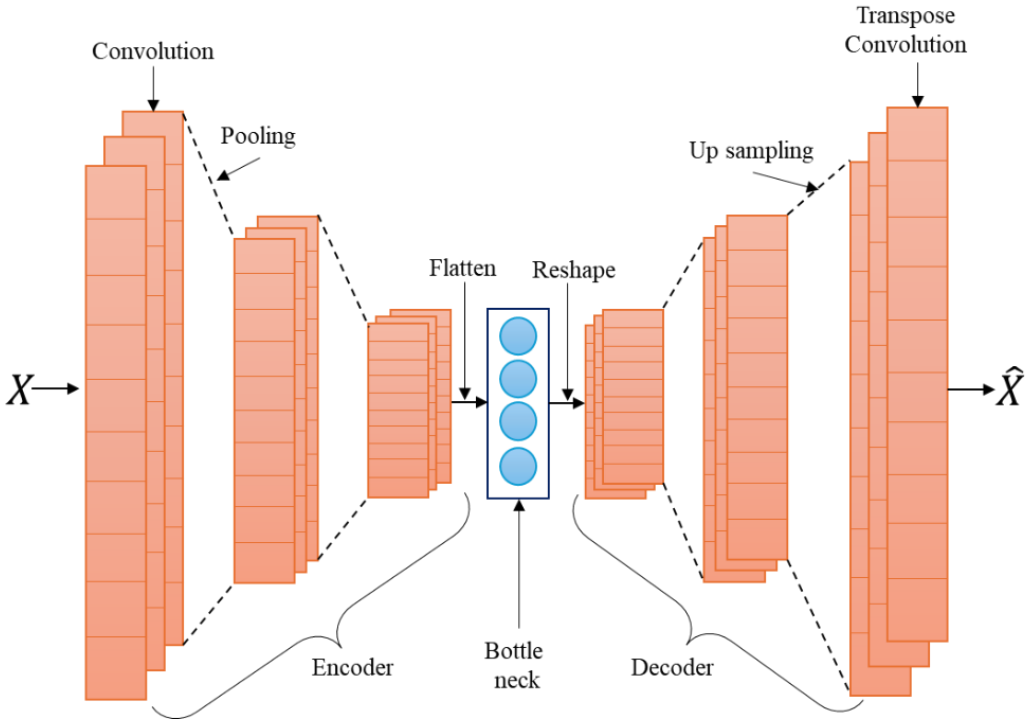


Figure 3.11: Convolutional autoencoder with fully connected bottle neck.

3.4 K-means algorithm

The K-means algorithm is a type of unsupervised machine learning, employed for clustering tasks. Its primary objective is to partition a dataset into K distinct clusters, where each data point belongs to the cluster with the nearest mean, serving as a prototype of the cluster. K-means is known for its simplicity, efficiency, and effectiveness in identifying underlying patterns within data. By minimizing the within-cluster variance, the algorithm aims to ensure that the data points within each cluster are as similar as possible, while clusters themselves are as distinct as possible. This versatility makes K-means a popular choice in various domains, including fault diagnosis, anomaly detection, among others. Despite its straightforward approach, K-means is powerful, offering valuable insights into the structure and relationships within complex datasets.

Given a set of observations $(x_1, x_2, \dots, x_n)$, where each observation is a d-dimensional real vector, K-means clustering seeks to divide the n observations into $k(\leq n)$ sets $S =$

$\{S_1, S_2, \ldots, S_k\}$ in such a way that the within-cluster sum of squares (WCSS), or variance, is minimized. The formal objective is to determine:

$$\arg\min \sum_{i=1}^{k} \sum_{x \epsilon S_i} \|x - \mu_i\|^2 = \arg\min \sum_{i=1}^{k} |S_i| \ \mathrm{Var}\, S_i \tag{3.15}$$

where μ_i is the mean or the centroid of points in S_i

$$\mu_i = \frac{1}{|S_i|} \sum_{x \epsilon S_i} x \tag{3.16}$$

$|S_i|$ is the size of S_i and $\|\cdot\|$ is the L^2 norm. This is equivalent to minimizing the squared differences between pairs of points within the same cluster:

$$\arg\min \sum_{i=1}^{k} \frac{1}{|S_i|} \sum_{x,y \epsilon S_i} \|x - y\|^2 \tag{3.17}$$

The K-means algorithm can be then summarized as follows. Given an initial set of k means $m_1, .., m_k$, the clusters are found alternating two steps:

Assign each observation to the cluster whose mean is closest, minimizing the squared Euclidean distance:

$$S_i{}^{(t)} = \{x_p : \left\|x_p - m_i{}^{(t)}\right\|^2 \leq \left\|x_p - m_j^{(t)}\right\|^2 \forall j, 1 \leq j \leq k\} \tag{3.18}$$

where each x_p is assigned to exactly one $S^{(t)}$.

The second step is to recalculate the means or centroids for the observations assigned to each cluster:

$$m_i^{(t+1)} = \frac{1}{\left|S_i^{(t)}\right|} \sum_{x_j \epsilon S_i^{(t)}} x_j \tag{3.19}$$

The algorithm converges when the WCSS stabilizes that means when the assignments no longer change.

4 PROPOSED DOMAIN ADAPTATION

Numerous works have sought to improve the generalization of machine learning models under variable working conditions by increasing their complexity and/or adding discrepancy terms to the loss function. These domain adaptation approaches perform well, but their application is limited to specific target domains that require retraining for a new working condition scenario, in addition to the high computational cost. Furthermore, some of these frameworks require label information from the source and target domains to perform domain synchronization, which in real industrial scenarios is unlikely to be available. In addition, some existing domain adaptation approaches that are fully unsupervised require prior information about the number of classes in each working condition.

In this context, a novel domain adaptation methodology for gearbox fault diagnosis under variable speed conditions is proposed. The introduced framework addressed the lack of labeled data by using unsupervised training without any prior health condition information. In addition, the domain adaptation is performed by finding correlations between data from source and target domain, avoiding data distribution assumptions.

The proposed method leverages the advantages of vibration analysis and deep learning. Knowledge about fault signatures is combined with automated feature extraction to improve the use of information and the fault diagnosis. In addition, a correlation metric is employed to perform the alignment of extracted features at different speed conditions.

This chapter is dedicated to explaining the proposed domain adaptation methodology. Figure 4.1 provides a flowchart of the proposed generalization approach which includes five phases: fault vibration signature calculation, band selection based on characteristic frequencies, data normalization, feature extraction by a convolutional autoencoder and finally feature synchronization through the correlation analysis.

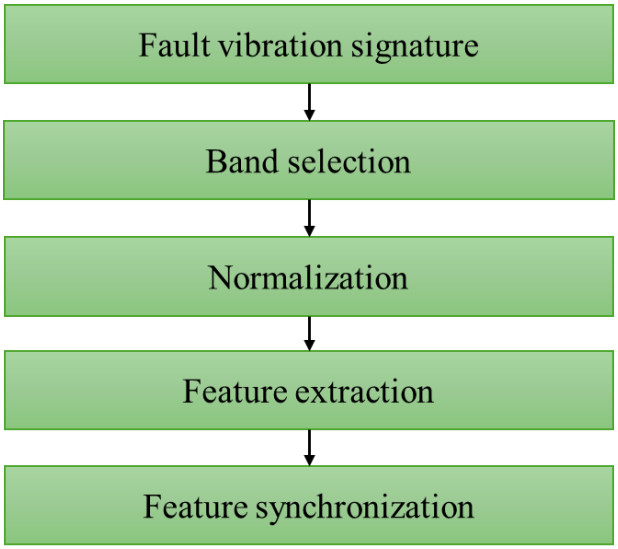


Figure 4.1: Flowchart of the proposed domain adaptation methodology.

4.1 Fault vibration signature

Works on fault diagnosis of rotating machinery using deep learning methods usually focus on the model complexity to achieve relevant prediction accuracy. These approaches rely on the model feature extraction power and extensive training for robustness of diagnosis. Most of them use the raw vibration signal as input to the neural network, seeking an end-to-end analysis that implies no data preparation or preprocessing. This results in the implementation of several regularization techniques to deal with the possible signal phase changes as well as the amplitude deviations caused by noise. On the other hand, several studies have shown that data preparation with signal processing techniques such as frequency spectrum (PANG et al., 2021), wavelet transform (HUANG et al., 2023), among others, improve accuracy and overall increase the prediction performance on test data. In addition, the number of parameters of the neural networks can be significantly reduced. In this context, taking advantage of signal processing, some techniques are applied to the raw vibration signals to obtain information about the fault vibration signatures.

The Fast Fourier Transform (FFT) is widely used in signal processing for machinery fault diagnosis. It is a powerful mathematical tool that transforms signals from the time domain to the frequency domain, providing valuable insights into the underlying frequencies present in

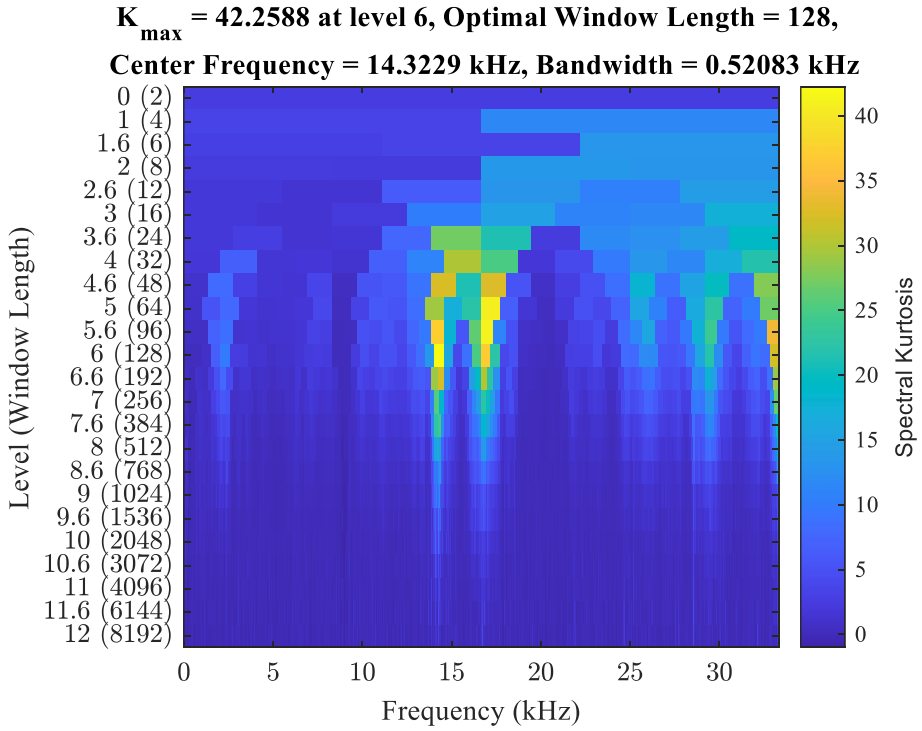
the signal. FFT helps in the identification of patterns associated with faults since different types of faults exhibit characteristic frequency signatures. For example, in rotating machinery, an unbalanced shaft might produce elevated vibration at the first harmonic of the rotational frequency. Thus, in this work, the Fast Fourier Transform is employed to extract the frequency content of the measured vibration and then identify relevant information about the faults.

Besides the Fast Fourier Transform, the proposed approach uses the envelope spectrum to obtain additional information of the vibration signatures. The envelope analysis is a stablished technique for identifying faults in rolling element bearings. For calculating the envelope spectrum, the vibration signal is filtered to improve the extraction of modulations caused by defects and minimize the interference from irrelevant components. The spectral kurtosis technique is applied to select the filtering band. Spectral kurtosis technique calculates kurtosis values locally across different frequency bands. This method is effective in identifying the frequency band with the highest kurtosis, or the highest signal-to-noise ratio. High spectral kurtosis values indicate high variance of power at the corresponding frequency, which makes spectral kurtosis a useful tool to locate nonstationary components of the signal. Once this specific frequency band is identified, a bandpass filter can be applied to the original signal. This filtering enhances the impulsive components of the signal, making it more suitable for envelope spectrum analysis. Figure 4.2 presents an example of spectral kurtosis for filter band selection to calculate the envelope spectrum and the corresponding envelope spectrum.

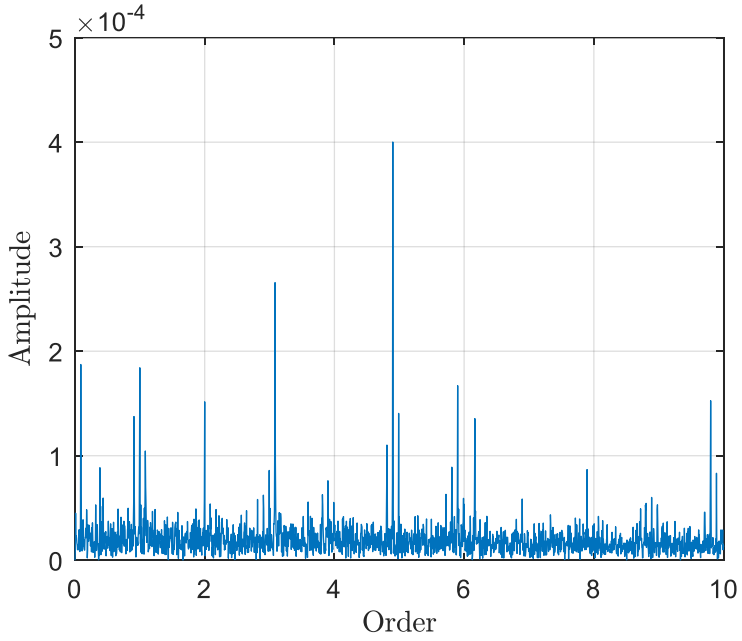
Analyzing machine vibration signature in the frequency domain presents some issues. Rotating machinery may be designed to work at different speeds. The same machine may work at different rotations depending on the operational conditions. For example, a gearbox employed in a wind turbine might operate at several rotational speeds depending on the wind speed. Another factor that may alter the rotational speed is load variation, producing engine slippage. Changes in working conditions influence the spectral content since most of the phenomena present in rotating machinery depend on the rotational speed.

To overcome the frequency shifting, in this work, the Fast Fourier Transform and the envelope spectrum are obtained in the order domain. For the case in which the system is operating in steady state but at different speed conditions, both spectra are easily calculated by normalizing the frequency content using the speed of the gearbox input shaft as reference. On the other hand, in the presence of fluctuating rotational speed conditions, resampling of the temporal vibration signal must be carried out using some order tracking technique, in order to obtain a stationary signal in the angle domain. In the last case, the signal from a tachometer or

an encoder must be available with the aim of computing the machine instantaneous rotational speed. Figure 4.3 provides the spectrum in frequency and order domains of a gearbox working at two different speeds.

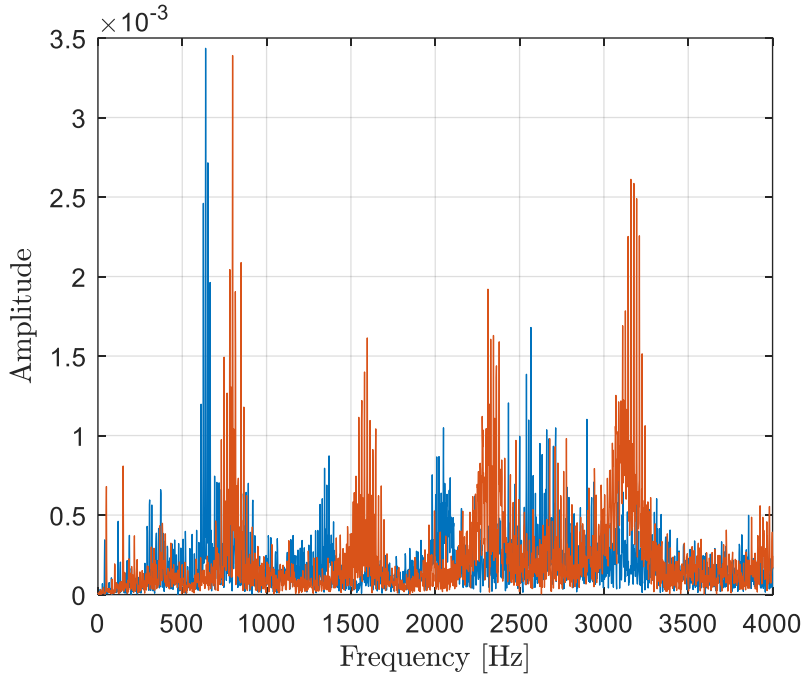


(a)

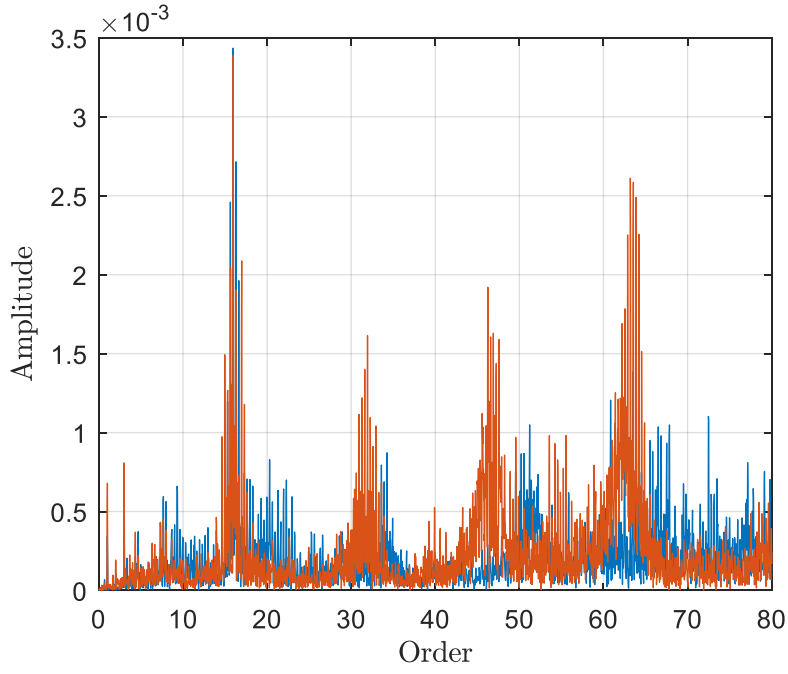


(b)

Figure 4.2: Signal filtering for calculation of envelope spectrum. (a) Spectral kurtosis method for band selection. (b) example of envelope spectrum.



(a)



(b)

Figure 4.3: Spectrum of a gearbox vibration signal from two different speed conditions. (a) Frequency domain. (b) Order domain.

The order analysis enables to align the fault characteristics frequencies, overcoming the shifting caused by speed variations. In this way, these frequencies can be compared and are suitable for the domain adaptation situation. Furthermore, order domain analysis facilitates the identification of the contributions of different machine components and consequently diagnosing faults more accurately.

Other authors have employed the order spectrum for domain adaptation as in the case of Ni et al., (2023) using a physics-informed residual network for diagnosing bearing faults under non-stationary conditions. Zheng et al., (2021) implemented a deep neural network that uses the order spectrum as priori knowledge for cross-domain diagnosis of rolling element bearings.

4.2 Band selection

Following the previous idea and seeking to make suitable use of the information obtained from the frequency and envelope spectra in the order domain, prior information about the failure signatures is applied to create the input vector that the convolutional autoencoder will use. From vibration analysis in fault diagnosis, it is known that each type of fault exhibits distinctive characteristic frequencies. With these frequencies, it is possible to identify patterns associated with specific flaws. Based on the previous information, this work combines knowledge from vibration analysis with deep learning to boost the feature extraction for fault classification.

The main idea here is to use bands corresponding to fault characteristic frequencies instead of the whole spectrum. For this purpose, based on the literature (JAMES I. TAYLOR, 2003), some bands are selected for fault diagnosis as described below:

Shaft issues are related to the amplitude increase of the first harmonics of the rotation frequency. For shaft imbalance, a high vibration amplitude is expected at 1x the shaft speed. Shaft misalignment can lead to an increased amplitude in the first three harmonics. A bent shaft is characterized by peaks at the first two harmonics of the shaft speed. Based on this information, the bands corresponding to the first four harmonics are chosen to identify shaft-related faults.

Elevated vibration at the gear mesh frequency (GMF) is related to gear problems. GMF is defined in eq. (4.1), where Z is the number of teeth of the gear and f_g is its rotational speed. Gear misalignment tends to excite the first three harmonics of the gear mesh frequency.

Eccentric gears generate elevated vibrations at 1x GMF and its harmonics. Additionally, if meshing gears have common factors (CF), peaks at GMF/CF and their harmonics can occur if one gear is eccentric. Backlash can cause excessive vibration at 2x GMF. Tooth defects are often associated with sidebands around the GMF. Broken, cracked, or chipped teeth may lead to increased sidebands around both the GMF and the gear's natural frequency. Excessive wear and clearance can also cause amplified sidebands around the GMF. Based on the previous information, bands corresponding to 1x GMF, 2x GMF, and 3x GMF for each gear pair are selected. Additionally, bands corresponding to GMF/CF and its harmonics as well as the first five side bands complete the set for gear fault diagnosis.

$$GMF = Z \times f_g \tag{4.1}$$

Bearing faults have characteristic frequencies well defined. Eq. (4.2) to (4.4) describe the ball pass frequency of the outer race (BPFO), the ball pass frequency of the inner race (BPFI), and the ball spin frequency (BSF), respectively. D_b is the rolling element diameter, D_p is the pitch diameter, N_r is the number of rolling elements for a single row, φ is the load angle and f_r is the rotational speed of the shaft.

$$BPFO = fr\frac{N_r}{2}\left(1 - \frac{D_b\cos(\varphi)}{D_p}\right) \tag{4.2}$$

$$BPFI = fr\frac{N_r}{2}\left(1 + \frac{D_b\cos(\varphi)}{D_p}\right) \tag{4.3}$$

$$BSF = fr\frac{D_p}{D_b}\left(1 - \frac{D_b^2\cos^2(\varphi)}{D_p^2}\right) \tag{4.4}$$

In addition, each characteristic frequency is related to a fault location: BPFO to the outer race, BPFI to the inner race, and BSF to rolling elements. Hence, bands corresponding to these frequencies are chosen from the envelope spectrum for diagnosing bearing faults. Table 4.1

summarizes the bands selected to diagnose faults in the gearbox. it is important to highlight that the bands are selected for every device component.

Table 4.1: Frequency bands for gearbox fault diagnosis

Gearbox component	Band
Shafts	$1 \times f_r$
	$2 \times f_r$
	$3 \times f_r$
	$4 \times f_r$
Gears	$1 \times GMF$
	$2 \times GMF$
	$3 \times GMF$
	$GMF \pm fr$
	GMF/CF
Bearings	$BPFO$
	$BPFI$
	BSF

Band selection has two main implications. The length of the input vector is considerably shorter, which results in fewer tunable parameter in the autoencoder. This allows the model to be built and trained using less computational power. In addition, the training time is significantly reduced, which allows for more runs in the testing phase. The second implication is the fact that the autoencoder has less information to process, and then the network extracts discriminative features more efficiently. This is because the convolutional autoencoder does not need to identify which data is important for classification, since the entire input vector contains relevant information.

4.3 Normalization

In machine learning applications is desired to normalize input data to make the training stable and avoid the exploding/vanishing gradient problem. For the case of computational vision, the most used normalization yields in dividing all the image pixels by the maximum

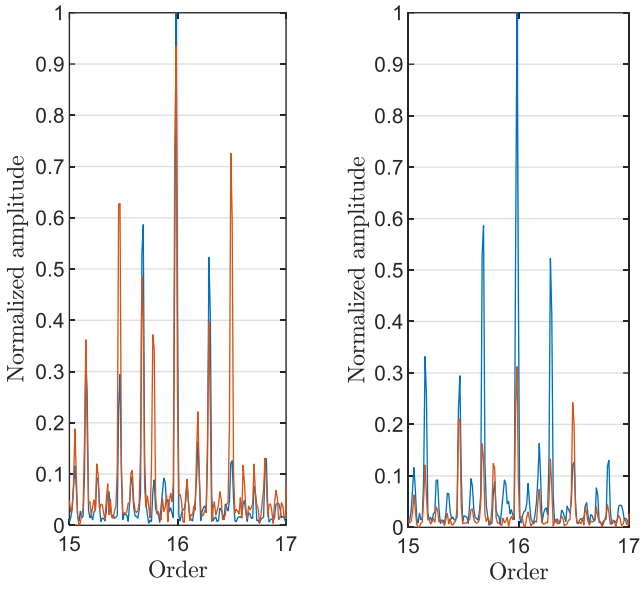
pixel value. For this application, the normalization works well because the values are always between 0 and 255. On the other hand, when data do not have a defined interval of values, the min-max mapping appears as the preferred option to normalize. This normalization implies that for each data vector, its maximum value will become 1 and the minimum zero. Eq. (4.5) presents its mathematical expression, where X' is the normalized vector, X is the input vector, X_{min} is the minimum value and X_{max} is the maximum value. In the field of fault diagnosis, this normalization is widely employed since feature vectors are usually vibration signals or images such as the time-frequency wavelet transform.

$$X' = \frac{X - X_{min}}{X_{max} - X_{min}} \tag{4.5}$$

Nevertheless, when analyzing the cited normalizations, some issues are detected in rotary machine fault diagnosis using bands selected from the order spectrum and envelope spectrum. As shown in the previous section, certain faults share the same characteristic frequency, with the primary differentiating factor often being the vibration amplitude. Consequently, normalizing each data vector individually can remove distinctions between fault signatures, potentially compromising diagnosis.

Considering the effects of normalizing each feature vector separately, this work normalize data globally for each speed condition. This means identifying the minimum and maximum values for all data associated with the same rotation speed, rather than for each individual vector, and then using these values as references for the min-max scaling. The global normalization ensures that differential features of fault signatures are not removed.

Figure 4.4 presents an example of the effects of normalization in a band related to a gear mesh frequency of two different health conditions. Figure 4.4 (a) corresponds to a conventional normalization and Figure 4.4 (b) to the proposed global normalization. As seen in the graph on the left, the difference in amplitude is removed and the convolutional autoencoder can then mistake the small variations for noise. In contrast, in the right graph, the discriminative characteristics are conserved.



(a) (b)

Figure 4.4: Effects of normalization in the discriminative characteristics. (a) Conventional normalization. (b) Global normalization.

4.4 Feature extraction

Feature extraction plays a crucial role in fault diagnosis since the extracted parameters are used for classification. In literature, strategies including statistical analysis and dimensionality reduction are implemented to obtain features that represent the analyzed data. However, machine learning approaches have gained visibility due to their performance and accuracy. Convolutional neural networks are one of the deep learning models extensively used for automated feature extraction.

In this work, a special type of convolutional neural network is employed for automated feature extraction: convolutional autoencoders. As described in section 3.3, autoencoders are non-supervised trained whose main objective is to reconstruct the input data from a compact representation. Autoencoders then extract parameters from data by themselves that synthesize the most important characteristics and encode them into a low-dimensional representation. These extracted features are invariant, which means that even when input data belonging to the same class present some variation due to noise or environmental conditions, the encoded data

follows the same pattern. Thus, invariant features are extracted from the input vector at different working speeds to perform the domain adaptation.

One advantage of using autoencoders instead of supervised convolutional neural networks is the flexibility in the values of the extracted features. Features extracted by a supervised neural network tend to be more rigid and defined into a well-defined range. On the other hand, features extracted by an autoencoder admit variation in their values allowing a greater generalization and then are suitable for domain adaptation.

4.5 Feature synchronization

Despite the features extracted by the autoencoder for data belonging to the same health condition in a specific working speed follow the same pattern, they are modified when domain changes, i.e., when the rotational speed of the gearbox varies. Therefore, these features are required to be aligned to diagnose faults in scenarios with variable working conditions.

Different approaches have been studied for features domain synchronization, some of them includes analyzing the data mutual information (CHEN, J. et al., 2021) , kullbac-Leibler divergence (KIM; LEE, 2023), entropy features and transfer learning (LI et al., 2021), among others. In this work, the correlation analysis (WU; LEE, 2011) is used to perform the features domain synchronization. The correlation between two feature vectors is defined as

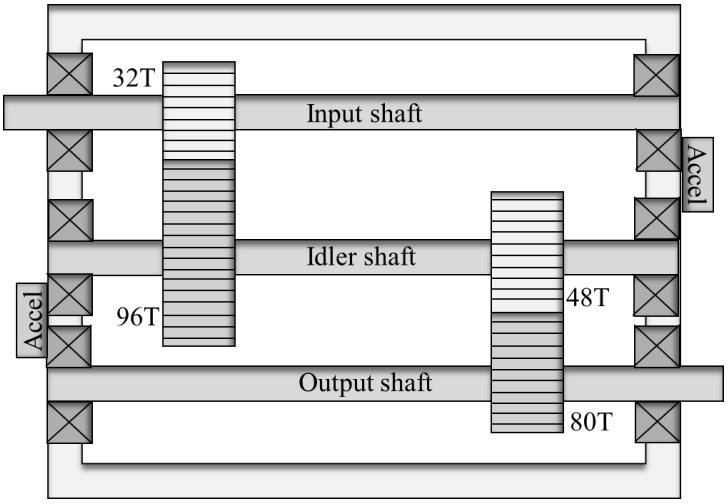
$$corr = \frac{FT_1 \cdot FT_2}{|FT_1| * |FT_2|} \tag{4.6}$$

where FT_1 and FT_2 are extracted feature vectors, $\cdot$ is the dot product and $|\cdot|$ is the largest singular value of a vector. The correlation value ranges from zero to one. The domain adaptation is then done by assigning the same class to the feature vectors with higher correlation.

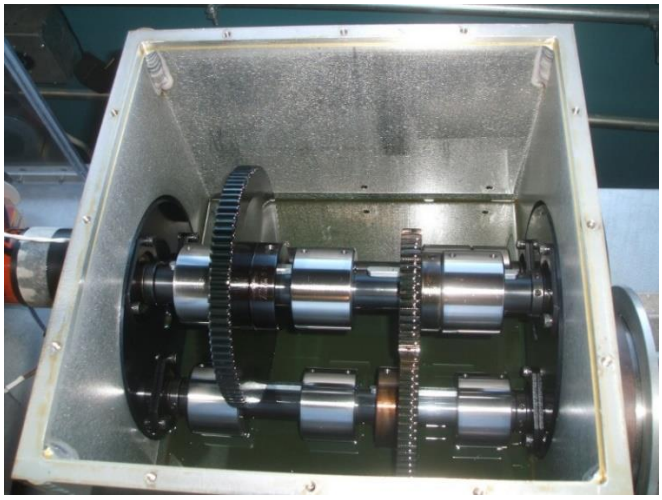
5 DATASET

The data to evaluate the proposed methodology come from the Prognostic and Health Management (PHM) Society 2009 competition. The dataset was measured from a representative general industrial gearbox whose schematic representation is shown is Figure 5.1. The gearbox has two stages arranged as follows:

- Input shaft: 32 teeth input pinion.
- Idler shaft: 96 teeth idler gear and 48 teeth idler pinion.
- Output shaft: 80 teeth output gear.



(a)



(b)

Figure 5.1: Generic industrial gearbox. (a) Schematic representation. (b) Test bench (PROGNOSTICS AND HEALTH MANAGEMENT SOCIETY, [s. d.])

From input to output shafts the gear reduction ration is $16/48 \times 24/40$, or 5 to 1 reduction. Shafts are supported by MB manufacturing ER-10K bearings whose specifications are presented in Table 5.1.

Table 5.1: Specifications of the ER-10K bearings.

Number of elements	8
Roller element diameter	$0.3125\ in$
Pitch diameter	$1.319\ in$
Contact angle	0

Considering the previous information, some characteristic frequencies of faults in gears and bearing are calculated as a function of the input shaft frequency, applying eqs. (4.1) to (4.4). The results are shown in Table 5.2.

Table 5.2: Fault characteristic frequencies in terms of the input shaft frequency.

Fault characteristic frequency	Value in terms of input shaft speed
GMF_1	$32f_{in}$
GMF_2	$16f_{in}$
$BPFO_{is}$	$3.05f_{in}$
$BPFI_{is}$	$4.94f_{in}$
BSF_{if}	$3.98f_{in}$
$BPFO_{ids}$	$1.01f_{in}$
$BPFI_{ids}$	$1.64f_{in}$
BSF_{ids}	$1.32f_{in}$
$BPFO_{os}$	$0.61f_{in}$
$BPFI_{os}$	$0.98f_{in}$
BSF_{of}	$0.79f_{in}$

The dataset consists of 6 different health conditions related to gears, bearings, and shaft problems. Table 5.3 provides a detailed description of the failure patterns. The data was collected under five input shaft speeds: $30, 35, 40, 45, 50\ Hz$. The vibration signals were measured by two Endevco 10 mv/g accelerometers with $\pm$ 1% error and resonance $> 45\ kHz$ located at input and output side of the gearbox. In addition, the rotation was measured by a tachometer using 10 pulse per revolution.

Table 5.3: Label pattern description of the gearbox

Label	Gear				Bearing						Shaft	
	32T	96T	48T	80T	IS:IS	ID:IS	OS:IS	IS:OS	ID:OS	OS:OS	Input	Output
$C1$	G	G	G	G	G	G	G	G	G	G	G	G
$C2$	C	G	E	G	G	G	G	G	G	G	G	G
$C3$	G	G	E	Br	B	G	G	G	G	G	G	G
$C4$	C	G	E	Br	In	B	O	G	G	G	G	G
$C5$	G	G	G	Br	In	B	O	G	G	G	Im	G
$C6$	G	G	G	G	G	B	O	G	G	G	Im	G

Legend: IS = Input shaft; ID=Idler shaft; OS = Output shaft; :IS = Input side; :OS = Output side. G= Good; C = Chipped; E=Eccentric; Br = Broken; B = Ball; In = Inner race; O = Outer race; Im= Imbalance.

The sampling frequency is $66.7\ kHz$ $(200/3\ kHz)$, the sampled time is 4 seconds, thus there are 266656 data points. To improve the usage of the dataset, the sliding window technique was implemented to create 40000 data points segments using a 1000 points stride for each fault condition.

This dataset has some particularities. As it was developed for a competition, the dynamics of the gearbox are intricate aiming to make difficult the fault diagnosis. The gear mesh frequencies share a relationship: the gear mesh frequency of the second stage is twice that of the first stage. Therefore, defects in gears that present an increase in the amplitude of the gear mesh harmonics represent a challenge to detect. In addition, some fault characteristic frequencies of bearings are pretty close to the shaft rotation and their harmonics. Nevertheless, the data set is composed by simultaneous faults occurring at different locations of the gearbox. this represents a more realistic situation of an industrial scenarios, where damages can be transferred to adjacent component generating compound faults.

The vibration signal and the corresponding frequency spectrum in the order domain of each class considered in this work are presented below. Figure 5.2 to 5.7 show data of class $C1$ to $C6$ respectively.

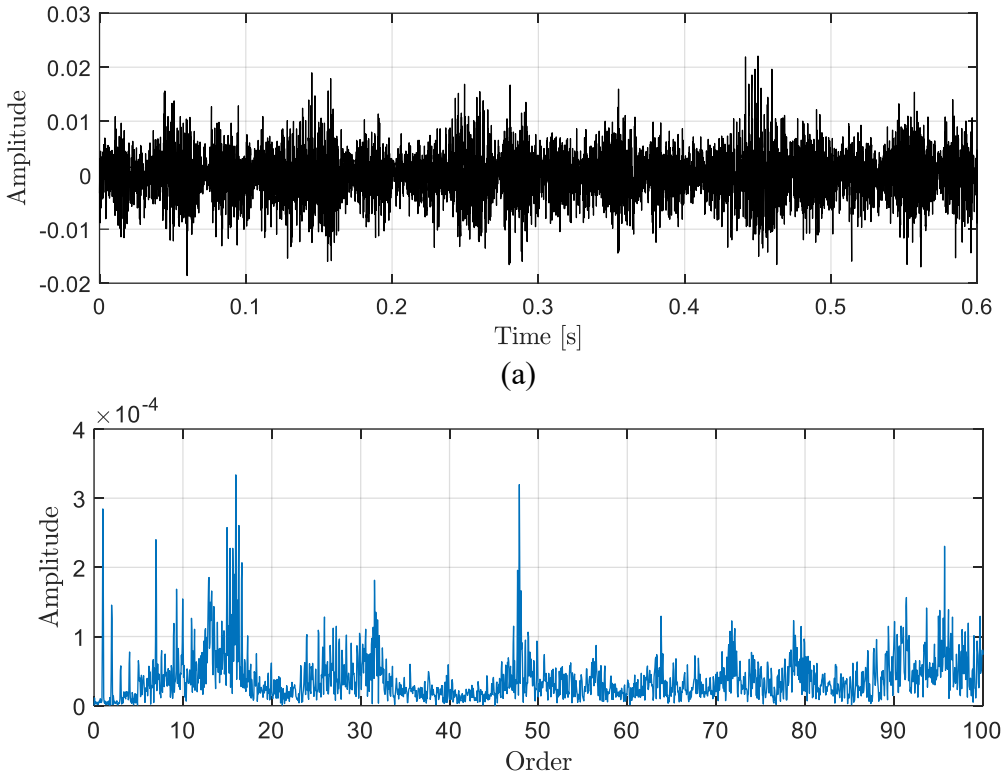


Figure 5.2: Sample of class $C1$. (a) Vibration signal. (b) Order spectrum.

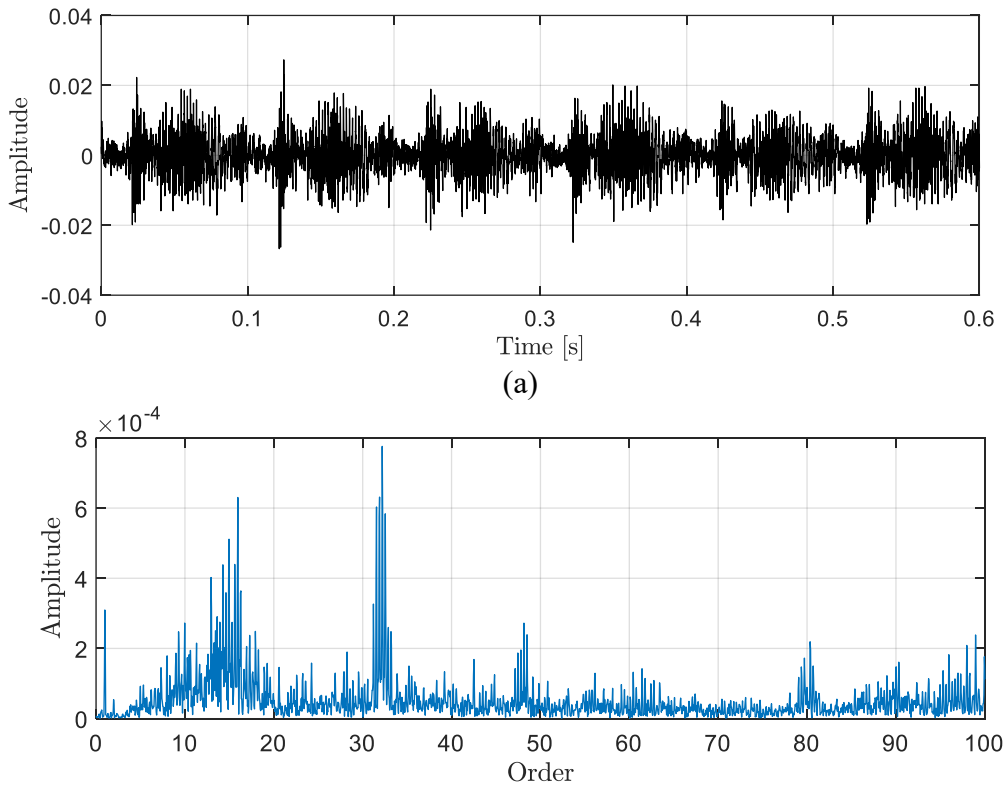
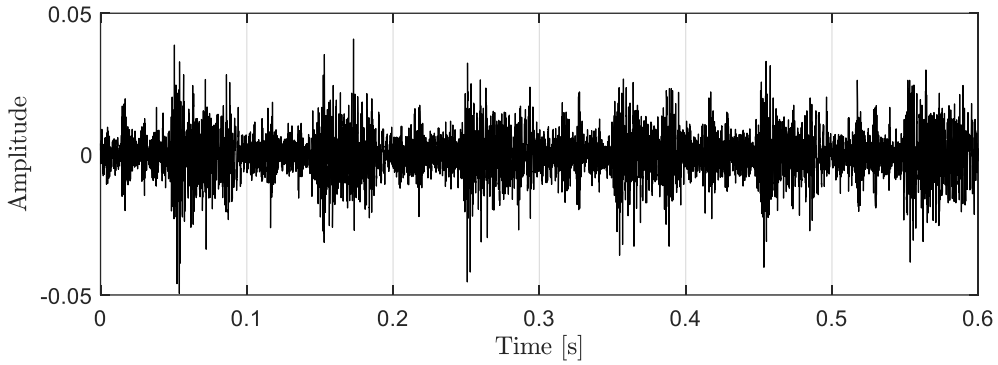
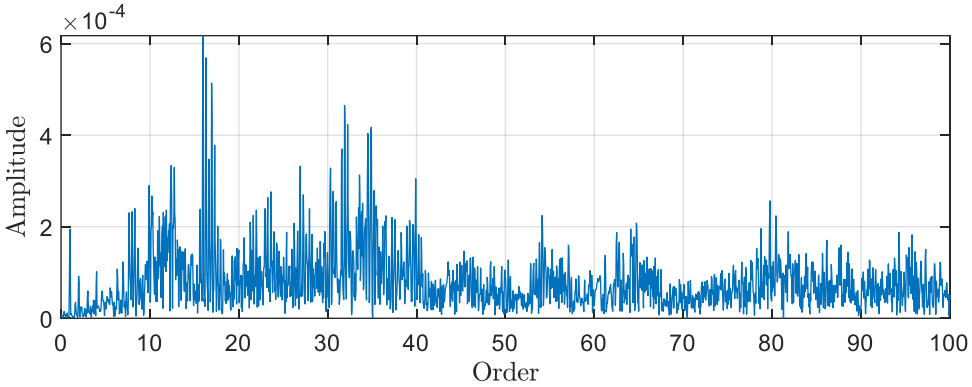


Figure 5.3: Sample of class $C2$. (a) Vibration signal. (b) Order spectrum.

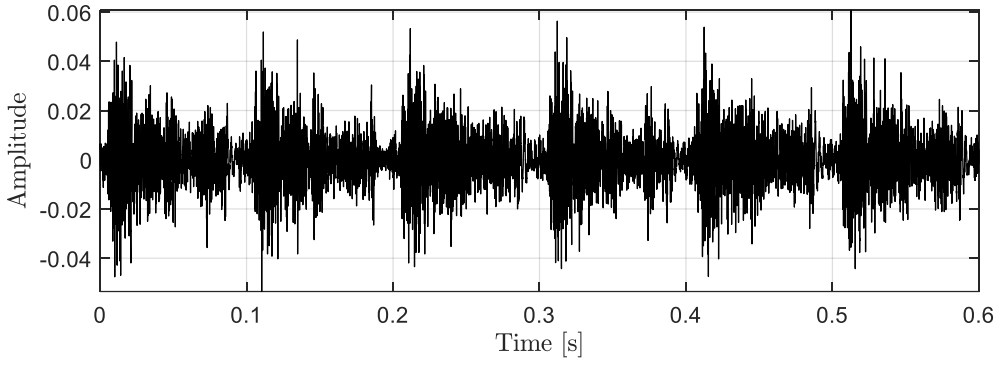


(a)

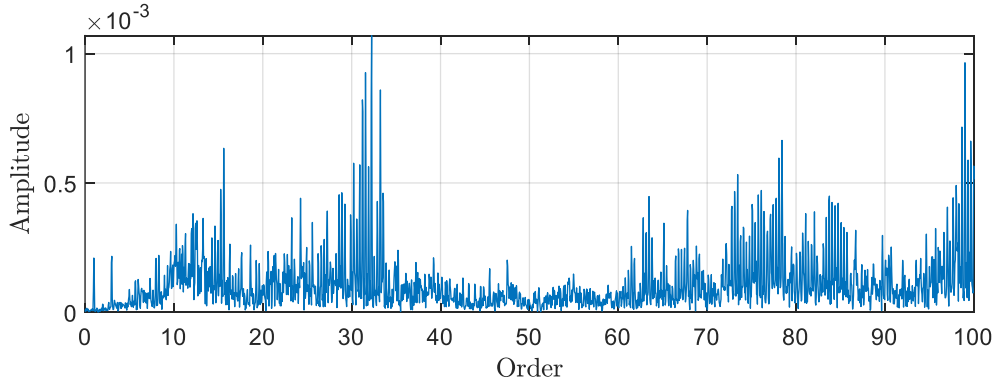


(b)

Figure 5.4: Sample of class $C3$. (a) Vibration signal. (b) Order spectrum.

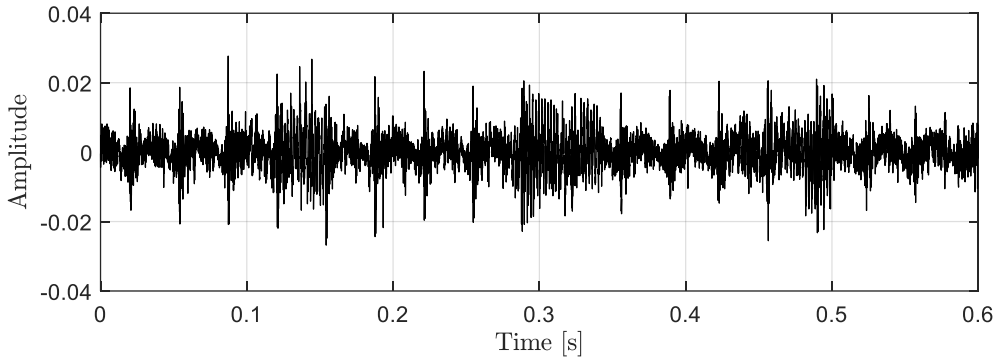


(a)

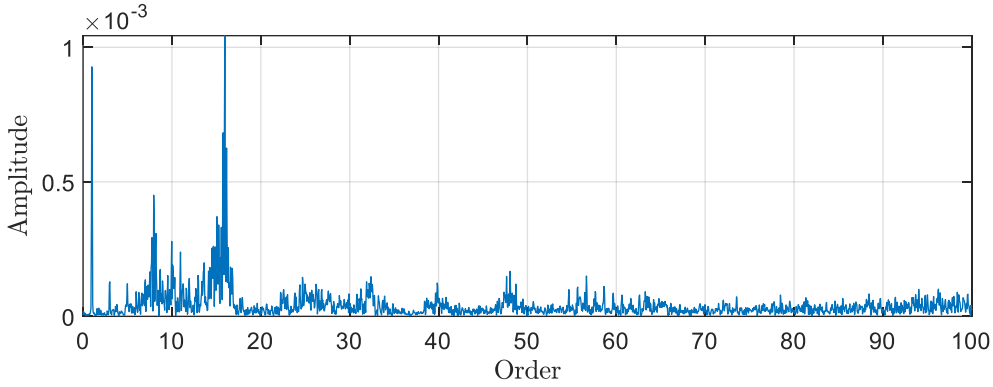


(b)

Figure 5.5: Sample of class $C4$. (a) Vibration signal. (b) Order spectrum.

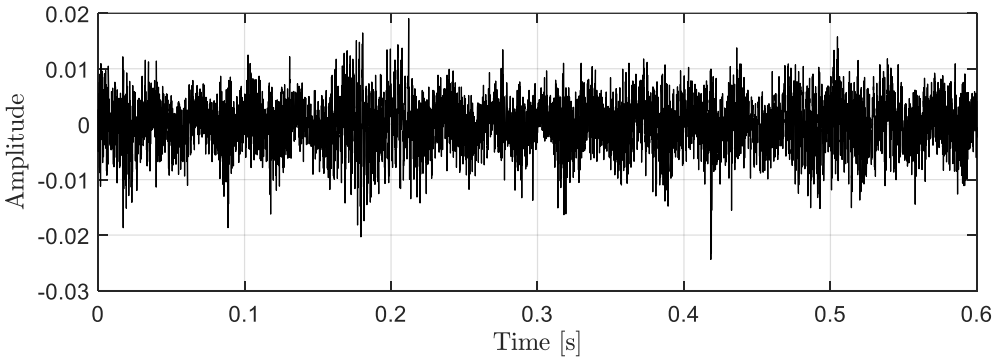


(a)

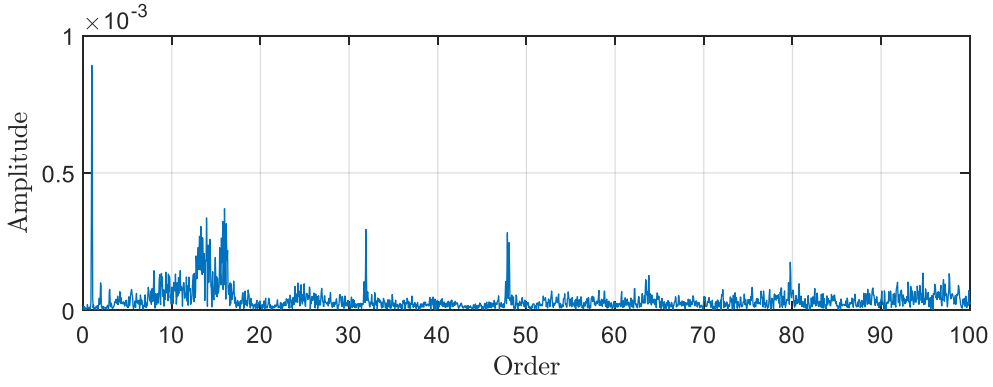


(b)

Figure 5.6: Sample of class $C5$. (a) Vibration signal. (b) Order spectrum.



(a)



(b)

Figure 5.7: Sample of class $C6$. (a) Vibration signal. (b) Order spectrum.

6 CLUSTERING-BASED GEARBOX FAULT DIAGNOSIS USING DOMAIN ADAPTATION

Rotating machinery fault diagnosis using deep learning methods have exponential grown in recent years. Research about the capabilities of these approaches have demonstrated their excellent performance in handling complicated mapping problems. However, some issues need to be solved. The training of the deep learning models usually requires large amount of labeled data which in real industrial scenarios are not available. In this context, unsupervised methods appear to overcome the lack of labeled data.

Studies on unsupervised fault diagnosis of rotating machinery has showcased outstanding results in detecting changes in the health state of the machine. For this purpose, strategies such as autoencoders and their variations identifies anomalies that allows to classify between a normal and faulty state. In addition, more current works also identifies several faults represented by different patterns. However, some limitations need to be addressed. The recognition of same type of faults across different operation conditions is still a challenge for unsupervised methods. This is because when operation conditions vary, the dynamic response of the machine is altered and consequently the extracted parameters undergo modifications.

Research on unsupervised domain adaptation has the objective of identifying similar patterns in the presence of domain shifts. Existing frameworks for unsupervised domain adaptation are semi-supervised since they demand label data from the source domain to perform the feature alignment. Therefore, they suggest some limitation in real applications where the machine health state is usually unknown. On the other hand, recent studies on fully unsupervised domain adaptation, i.e., no label data from the source and target data is required, but prior information about the numbers of classes in each domain is necessary to perform the domain adaptation.

In this scenario, this work proposes a novel clustering-based domain adaptation methodology for unsupervised gearbox fault diagnosis under variable speed conditions. The methodology and results are presented below.

6.1 Proposed methodology

One common approach in unsupervised fault diagnosis is clustering, where data points are grouped based on their similarity. For this study case, the domain adaptation is based on the clustering of the data belonging to the same class in each domain. Figure 6.1 presents a flow chart of the methodology used to cluster and classify different gearbox faults across different speed conditions without using labels.

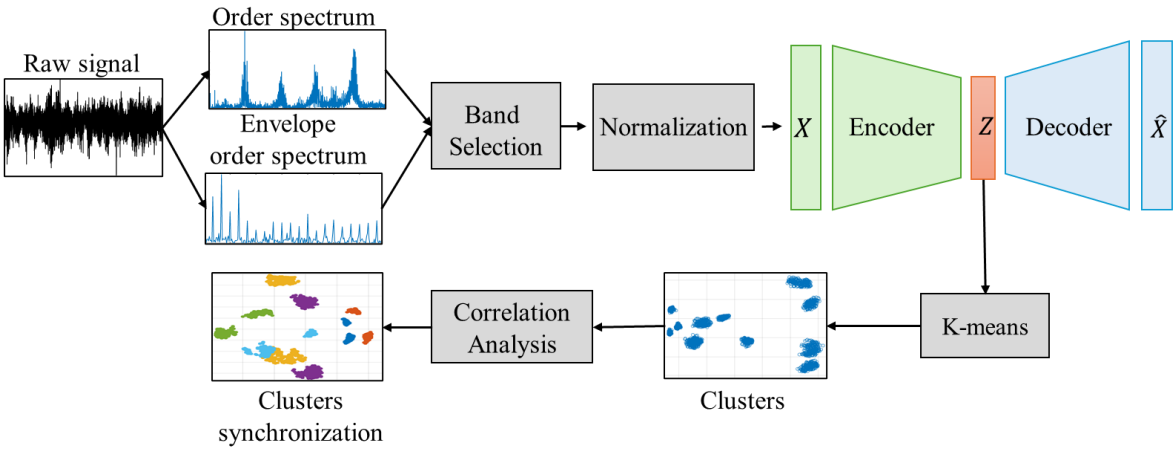


Figure 6.1: Proposed framework for clustering-based gearbox fault diagnosis using domain adaptation.

The framework employs the domain adaptation method explained in Chapter 4 and the K-means algorithm. First, from the raw vibration signals, the frequency and envelope spectrum are calculated in the order domain to obtain the fault signature. After this, bands are selected according to the representative characteristic frequencies to construct the vector that will serve as input for the autoencoder. This vector is then normalized applying the global normalization according to each domain and processed by the autoencoder to reduce its dimension and extract discriminative features. As the main objective here is to cluster, the K-means algorithm is employed to group the features and classify the different classes in each domain. Finally, the clusters belonging to the same class in different gearbox speeds are identified by executing the correlation analysis.

The architecture of the employed autoencoder is as follows: the encoder comprises three convolutional layers with 12, 24 and 48 filters respectively, where the kernel size is 12. Among the convolutional layers, the max poling layers perform down sampling with a pool size of 2. All the layers have a stride of 2. The latent space represented by the bottleneck is a fully connected layer of 80 dimensions. The activation function of the convolutional layers is ReLU while for the fully connected layers is sigmoid. The decoder has a similar configuration. Table 6.1 presents the complete architecture of the autoencoder. The preparation of the input vector that includes the calculation of the frequency and envelope spectra and the band selection was done using MATLAB. Furthermore, the convolutional autoencoder was constructed and trained in Python environment using Keras that is a high-level application programming interface developed by Google for implementing neural networks.

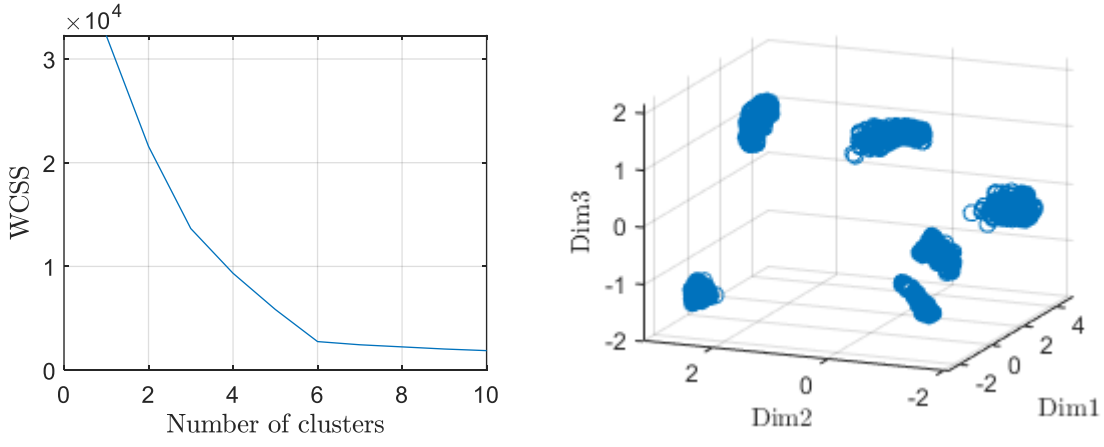
Table 6.1: Architecture of the employed convolutional autoencoder.

Layer type	Output shape	Number of trainable parameters
Convolutional	(486, 12)	456
Max Pooling	(243, 12)	0
Convolutional	(116, 24)	3480
Max Pooling	(58, 24)	0
Convolutional	(24, 48)	13872
Max Pooling	(12, 48)	0
Flatten	576	0
Dense	80	46160
Dense	576	46656
Reshape	(12, 48)	0
Up sampling	(24,48)	0
Convolutional transpose	(58, 48)	27696
Up sampling	(116, 48)	0
Convolutional transpose	(243, 24)	15000
Up sampling	(486, 24)	0
Convolutional transpose	(982, 12)	3486
Convolutional	(982, 1)	145

6.2 Clustering results

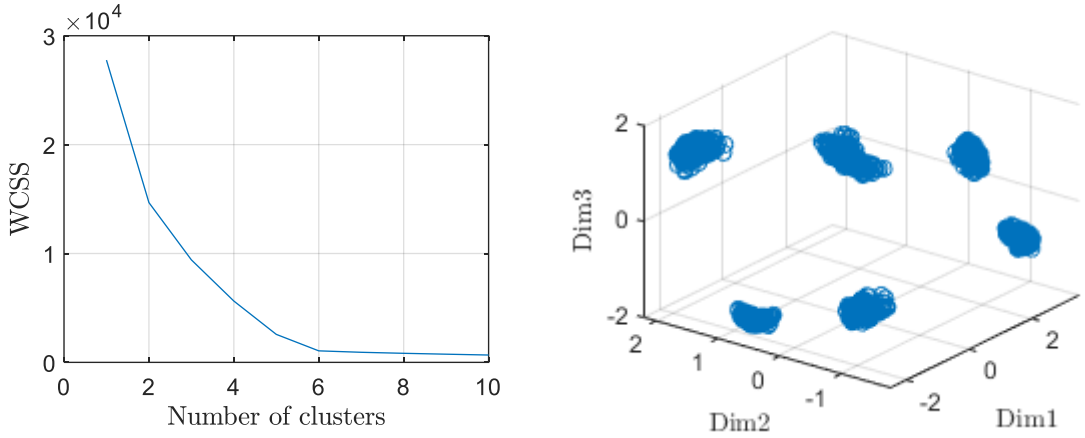
The first analysis corresponds to the clustering results in each gearbox speed separately. Here the goal is to correctly identify the number of classes for each working condition. To evaluate the results two approaches are employed: the cluster visualization and the elbow method. The feature vectors provided by the autoencoder has usually dimensions greater than 3, and therefore, it is not possible to graph them. Thus, the principal component analysis (PCA) is used to reduce the dimension of the feature vector to 3 aiming at their visualization. On the other hand, the elbow graph shows the within-cluster sum of square (WCSS) values on the y-axis corresponding to the possible number of classes. The optimal number of clusters is the point at which the line forms an elbow, which means that from this point the variation of WCSS value does not present an abrupt decrease.

Figure 6.2 and 6.3 presents the WCSS plots and the clusters visualization for data measured at 30 and 45 Hz respectively containing six fault types. In both cases, the proposed method achieved excellent clustering effect. Sample belonging to the same class are drawn together and different type of fault are well separated. In addition, when analyzing the elbow method, for both rotation speeds, the point of optimal number of clusters is clearly defined corroborating the results obtained by visualization. Accurate identification of the number of classes and their respective clusters in each domain is essential for successful domain adaptation at different speed rotations.



(a) (b)

Figure 6.2: Clustering results at 30 Hz. (a) Elbow plot. (b) Clusters visualization.



(a) (b)

Figure 6.3: Clustering results at 45 Hz. (a) Elbow plot. (b) Clusters visualization.

6.3 Domain adaptation results

Once the number of classes in each domain is defined, the next step is to identify which clusters represent the same type of faults under different working condition. For this purpose, the identification of clusters belonging to the same health condition is performed through the correlation analysis explained in section 4.5. A correlation matrix is obtained by calculating the correlation between the cluster in the source and target domain. The columns represent classes from the source domain while the rows represent classes from the target domain. Some tests are performed to evaluate the effectiveness of the proposed approach.

Tables 6.2 to 6.4 provide the correlations matrixes between the clusters for scenarios where the operation speeds are considered close. The greatest correlations are highlighted. For the three cases, clusters belonging to the same fault type present the greatest correlation, indicating that the proposed domain adaptation has excellent performance. Correlations between clusters at 30 and 35 Hz are slightly lower compared to the other two cases. This can be explained due to the influence of the machine speed on the vibration signature. Phenomena such as shaft unbalancing and gear eccentricity are strongly associated with the increase of amplitude regarding the rotational frequency. For this study case, amplitude in bands related to shaft and gears issues grows at higher speeds and therefore the vibration signature is more evident in comparison to lower frequencies. Thus, the extracted features reflect these effects.

Table 6.2: Correlation matrix between 30 and 35 Hz.

	$C1$	$C2$	$C3$	$C4$	$C5$	$C6$
$C1$	$\mathbf{0.8667}$	0.6862	0.4569	0.4656	0.5528	0.5609
$C2$	0.7253	$\mathbf{0.7937}$	0.5783	0.7143	0.4310	0.5865
$C3$	0.4725	0.5418	$\mathbf{0.7935}$	0.6920	0.5073	0.4142
$C4$	0.4135	0.6212	0.7287	$\mathbf{0.7503}$	0.5544	0.4516
$C5$	0.5608	0.5085	0.6993	0.4825	$\mathbf{0.8497}$	0.6814
$C6$	0.6761	0.5442	0.5677	0.5034	0.6810	$\mathbf{0.7277}$

Table 6.3: Correlation matrix between 35 and 40 Hz.

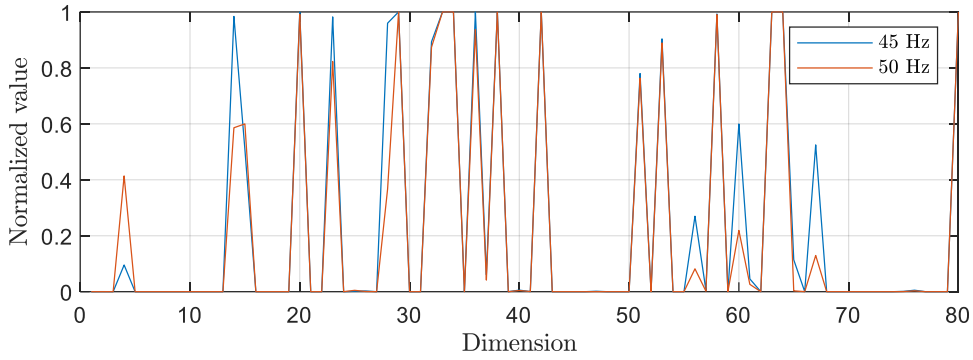
	$C1$	$C2$	$C3$	$C4$	$C5$	$C6$
$C1$	$\mathbf{0.9010}$	0.5256	0.4388	0.3770	0.4780	0.6058
$C2$	0.6799	$\mathbf{0.8005}$	0.5754	0.5789	0.5647	0.5797
$C3$	0.5392	0.6556	$\mathbf{0.7840}$	0.6977	0.6801	0.5788
$C4$	0.5183	0.6097	0.6966	$\mathbf{0.7921}$	0.5700	0.5195
$C5$	0.6412	0.3291	0.5221	0.3753	$\mathbf{0.7875}$	0.6198
$C6$	0.7532	0.3718	0.4957	0.3608	0.5982	$\mathbf{0.8523}$

Table 6.4: Correlation matrix between 45 and 50 Hz.

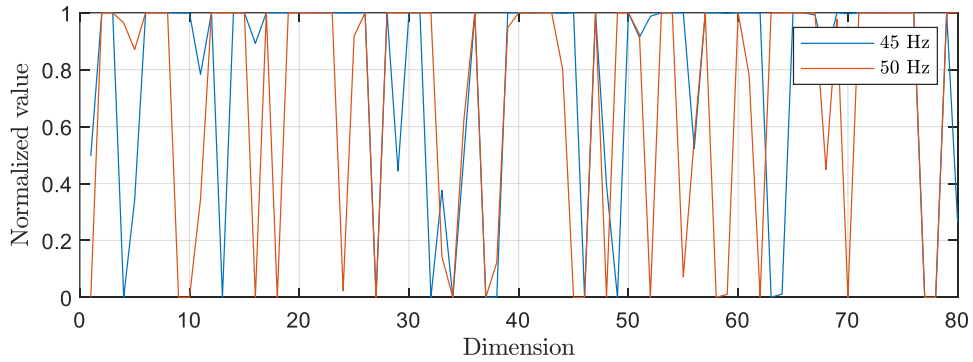
	$C1$	$C2$	$C3$	$C4$	$C5$	$C6$
$C1$	$\mathbf{0.9153}$	0.6893	0.4820	0.4789	0.6004	0.6746
$C2$	0.7306	$\mathbf{0.8468}$	0.6133	0.6658	0.6689	0.6604
$C3$	0.5684	0.6510	$\mathbf{0.7881}$	0.7137	0.7322	0.6116
$C4$	0.5224	0.5896	0.7163	$\mathbf{0.7830}$	0.6327	0.5660
$C5$	0.6998	0.4535	0.5384	0.4223	$\mathbf{0.8601}$	0.6770
$C6$	0.7995	0.5173	0.5545	0.4547	0.6900	$\mathbf{0.8462}$

Figure 6.4 depicts the feature vectors of two health conditions at two different speeds. Figure 6.4 (a) a corresponds to class $C1$ and Figure 6.4 (b) to class $C3$. As shown in this figure, most of the values of the feature vector in both cases are similar, demonstrating high correlation. However, for the class $C3$, some values of the features vectors are different, impacting directly the correlation coefficient as shown in Table 6.4. Another factor to analyze is how the classes

can be differentiated comparing the extracted feature vectors. The class $C1$ has most of the values close to zero and some little positions of the vectors close to one. In contrast, the values of the feature vectors of the class $C3$ are more distributed between zero and one.



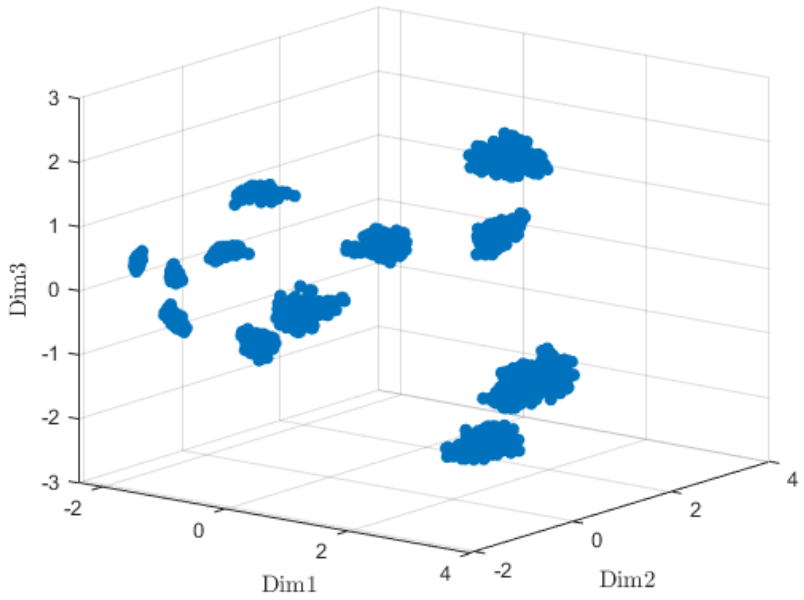
(a)



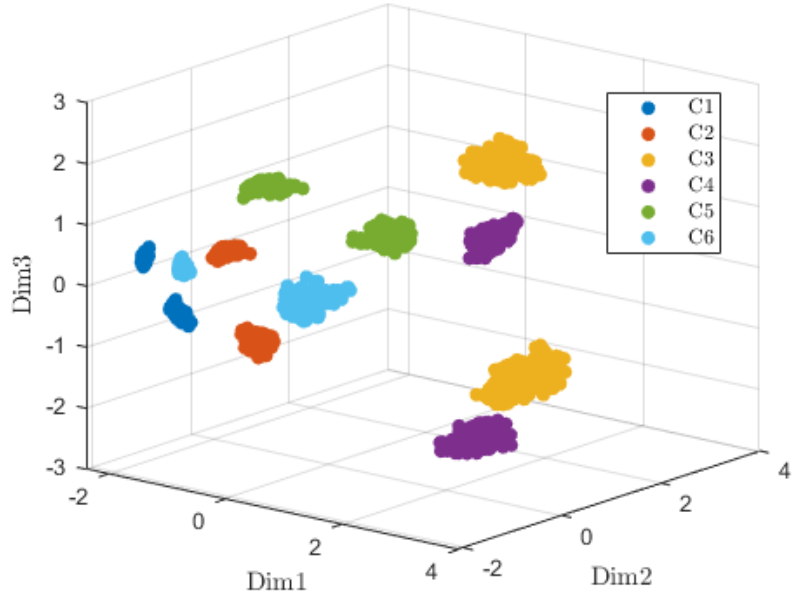
(b)

Figure 6.4: Extracted features for two different speed conditions. (a) Class $C1$. (b) Class $C3$.

Figure 6.5 shows the clusters corresponding to the six different health conditions at two different speeds after the correlation analysis. The identification of clusters representing data from the same fault is not intuitive. One can suggest, based on visualization, that the closest clusters belong to the same class. However, as seen in Figure 6.5 (b), only clusters of class C1 are relatively close. Clusters of other classes are scattered in the space and closer to different fault types instead of the same health condition.



(a)



(b)

Figure 6.5: Cluster from two different speed conditions. (a) Before correlation analysis. (b) After correlation analysis.

An analysis was also carried out incorporating data from more distant speeds. Table 6.5 displays the correlation between clusters at $50\,Hz$ and $40\,Hz$, while Table 6.6 presents correlations between $50\,Hz$ and $30\,Hz$. In the former case, clusters representing identical faults

were correctly identified, albeit with decreased correlation values compared to previous tests. This decrease is anticipated due to the impact of rotational speed on fault signatures. However, a misclassification occurred when clustering data from 50 Hz and 30 Hz. Specifically, fault label $C2$ exhibited a higher correlation with label $C4$ instead of its own. This anomaly can be attributed to the similarity between the signatures of patterns $C2$ and $C4$ at 30 Hz. Both labels correspond to a chipped tooth on the 32-tooth gear and eccentricity on the 48-tooth gear. Although $C4$ also involves a broken tooth on the 80-tooth gear, this occurs at the same gear mesh frequency as the eccentricity, which is a dominant feature at this frequency. Alongside the similarity in fault signatures, another factor contributing to the misclassification was the significant difference in gearbox speeds, leading to notable changes in the amplitude of the characteristic frequencies.

Table 6.5: Correlation matrix between 40 and 50 Hz.

	$C1$	$C2$	$C3$	$C4$	$C5$	$C6$
$C1$	**0.7308**	0.4244	0.2390	0.1668	0.5467	0.5395
$C2$	0.5806	**0.6252**	0.4764	0.4025	0.5019	0.4519
$C3$	0.4222	0.5456	**0.7387**	0.6342	0.4518	0.3282
$C4$	0.4922	0.4896	0.5307	**0.6650**	0.3809	0.3558
$C5$	0.5423	0.4115	0.2305	0.1380	**0.5763**	0.4390
$C6$	0.5279	0.4617	0.2962	0.2406	0.5598	**0.6024**

Table 6.6: Correlation matrix between 30 and 50 Hz.

	$C1$	$C2$	$C3$	$C4$	$C5$	$C6$
$C1$	**0.6310**	0.6033	0.5453	0.5157	0.4760	0.5052
$C2$	0.5170	0.6292	0.5804	**0.6329**	0.4999	0.4587
$C3$	0.5353	0.6316	**0.6804**	0.4464	0.5021	0.5204
$C4$	0.4716	0.5874	0.5300	**0.6825**	0.4394	0.4009
$C5$	0.4442	0.4064	0.2960	0.2776	**0.5973**	0.4918
$C6$	0.5102	0.4561	0.4189	0.3782	0.5285	**0.5492**

The influence of the normalization in fault identification among domain shifts was also investigated. For this purpose, two tests using data from close speeds were individually

normalized and then the correlation analysis was performed to evaluate the effects on domain adaptation. Table 6.7 and 6.8 presents the results. As observed in the matrixes, several faults are misclassified in both cases and the overall correlation values are lower than when the proposed normalization is employed as shown in Tables 6.2 to 6.4. These poorer outcomes are mainly due to the removal of the amplitude changes in faults associated with the same characteristic frequencies. Besides, for these two tests, the autoencoder training was considerably slower and the mean squared error values were not significantly improved across numerous iterations. The extraction features process is consequently compromised and affects the correlation analysis.

Table 6.7: Correlation matrix between 30 and 35 Hz using conventional normalization.

	$C1$	$C2$	$C3$	$C4$	$C5$	$C6$
$C1$	$\mathbf{0.6672}$	0.5133	0.5625	0.4929	0.2919	0.3172
$C2$	0.6039	0.4840	0.6586	$\mathbf{0.6894}$	0.3412	0.3353
$C3$	0.5673	0.5704	$\mathbf{0.7442}$	0.6690	0.2722	0.1859
$C4$	0.5649	0.5793	$\mathbf{0.7158}$	0.6987	0.3106	0.2279
$C5$	0.5610	0.5662	0.4493	0.4384	$\mathbf{0.6479}$	0.5194
$C6$	$\mathbf{0.6222}$	0.4887	0.3138	0.3946	0.3789	0.4773

Table 6.8: Correlation matrix between 45 and 50 Hz using conventional normalization.

	$C1$	$C2$	$C3$	$C4$	$C5$	$C6$
$C1$	$\mathbf{0.6966}$	0.4411	0.5235	0.4333	0.2649	0.5467
$C2$	0.4520	0.6647	0.6279	0.6632	$\mathbf{0.6705}$	0.3986
$C3$	0.2719	0.3957	$\mathbf{0.5466}$	0.5208	0.4588	0.1486
$C4$	0.3263	$\mathbf{0.6847}$	0.4639	0.4564	0.3762	0.1677
$C5$	0.5454	0.6004	0.5244	0.5116	$\mathbf{0.6871}$	0.6133
$C6$	$\mathbf{0.6865}$	0.3679	0.4330	0.2778	0.1337	0.6027

The impact of the number of dimensions of the extracted features over the domain adaptation was assessed. Data from two adjacent gearbox speeds were processed by the autoencoder varying the size of the bottleneck. Figure 6.6 shows the number of classes correctly identified as a function of the encoded dimension. The correlation analysis improved as the number of dimensions increased. This analysis is important since the size of the

bottleneck does not highly influence the identification of the number of classes in each gearbox rotation speed. However, as seen in this figure, considering few dimensions for the extracted features may lead to misclassifications across different speed conditions.

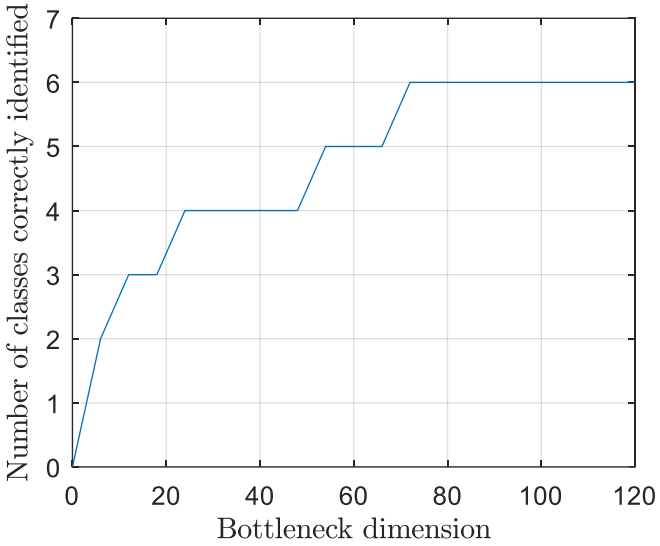


Figure 6.6: Relation between the number of dimensions in the latent space and the number of classes correctly identified.

Finally, the presented methodology achieved prominent results in diagnosing gearbox faults in an unsupervised manner. It is worth noting that works that use domain adaptation schemes as adversarial training and discrepancy metrics, among others, are not totally unsupervised, since they use prior label information from the source domain to perform the feature alignment. In contrast, the described approach is completely non-supervised which implies that not require any label information. This allows the implementation of the strategy in a variety of datasets and helps to overcome the lack of labeled data for machine learning training. Furthermore, the proposed framework enables the identification of the number of classes in each domain, addressing limitations of fully unsupervised domain adaptation strategies. On the other hand, the computational complexity of the algorithms found in the literature suggests some difficulty in their application in real industrial situations. In contrast, the proposed approach requires low computational power, and its implementation is simplified compared to other frameworks.

7 RECONSTRUCTION-BASED GEARBOX FAULT DIAGNOSIS USING DOMAIN ADAPTATION

Supervised training is the most exploited application of deep learning methods for rotating machinery fault diagnosis. Since the emergence of the first algorithms to the present day, extensive research has studied the scope of these approaches, demonstrating versatility and high performance in comparison to conventional methods. Most of the works on the area focused in analyzing how diverse structures of neural networks provides betters results in modelling complex mapping problems. Algorithms such as convolutional and recurrent neural networks are in trend because of their high capability of automatically extracting features and handling time-series data, respectively.

Generalization of the models — the performance of the neural networks using data different to the one used to train —has been a constant concern for researchers. Along the years, techniques have been developed to improve the accuracy over test data. Approaches such as L1 and L2 regularization enhance the generalization of the models. Dropout also appears as a strategy to ovoid overfitting. However, these methodologies work well when test data belong to the same distribution as those used to train the model.

When deep learning methods are tested using data from a different domain that the one which the model is trained, their performance usually decreases significantly. For example, a model to predict faults in a rotating machinery was trained using data measured at three different speeds. When the model is tested using new data measured at the same three speeds, the accuracy of the results is good. However, whether the model is tested using data from a fourth speed, the accuracy of the diagnosis is compromised.

Domain adaptation methodologies emerged as strategies to improve the accuracy of deep learning models when data domain shifts. Approaches such as the maximum mean discrepancy and adversarial training gained popularity in this field. The issue with these frameworks is that domain adaptation is usually performed to a specific target domain, demanding retraining for a new working condition. A third alternative is reconstruction methods. They are more flexible than the previously mentioned methods and their application has grown due to the lower computational power required.

In this context, this work proposes a new reconstruction-based domain adaptation methodology for gearbox fault diagnosis under variable speed conditions. The main idea is to train the reconstruction model using data measured at specific speed conditions and test it using data measured at speeds different from those used in training. The proposed methodology and results are shown below.

7.1 Proposed methodology

Figure 7.1 presents the flowchart employed for gearbox fault diagnosis using the proposed domain adaptation. From raw signal, order and envelope spectra are calculated to obtain the vibration signatures. Bands of interest are selected to construct the vector that serves as input for the autoencoder. This vector is normalized and then processed by the network to extract invariant features. The correlation analysis is performed and finally a classifier assigns the labels corresponding to each fault pattern. The architecture of the employed autoencoder is the same as the used in section 6.1, presented in Table 6.1.

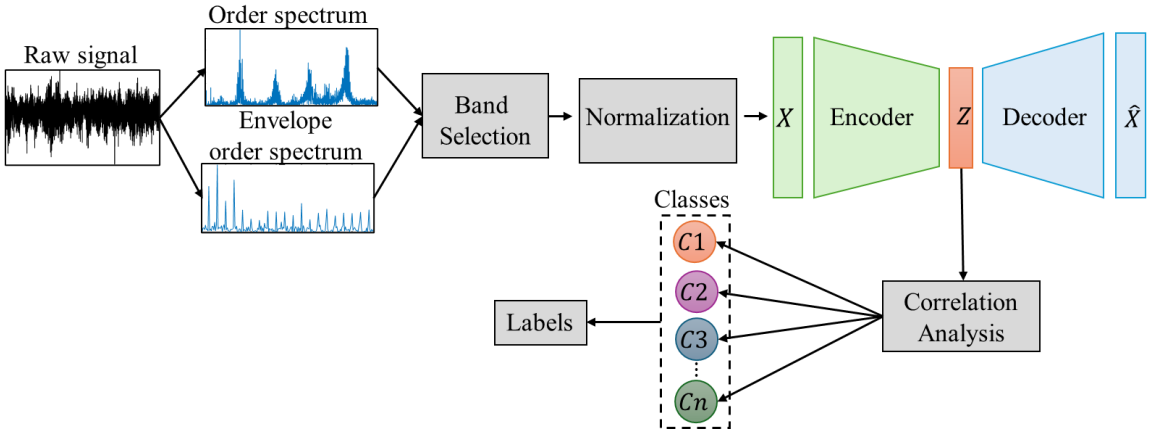


Figure 7.1: Proposed framework for gearbox fault diagnosis using a reconstruction-based domain adaptation.

To compare the performance of the proposed domain adaptation, two convolutional neural network models are built. One of the models is fed with the spectrum in the order domain while the other use as input bands selected from the frequency and envelope spectra in the order

domain. Figures 7.2 and 7.3 show a schematic representation of the convolutional neural network models. In addition, Table 7.1 presents the architecture of the CNNs.

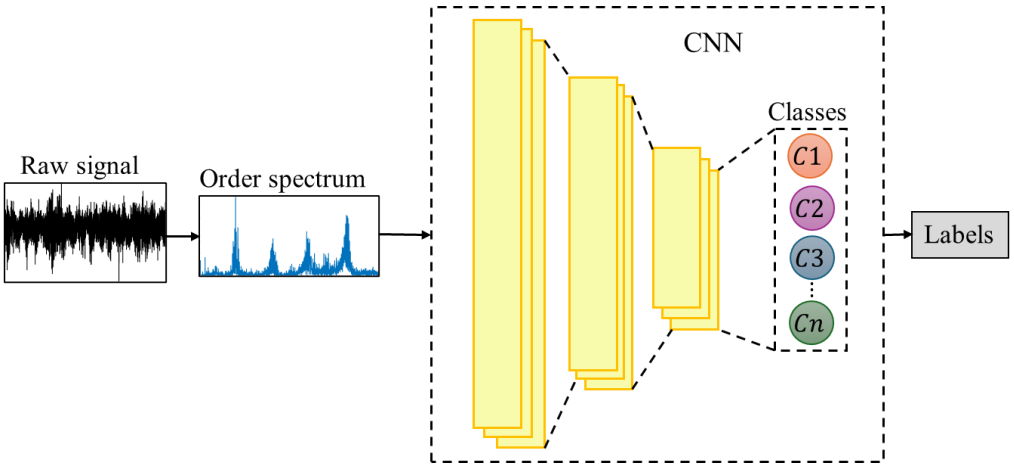


Figure 7.2: Convolutional neural network model using the spectrum in the order domain.

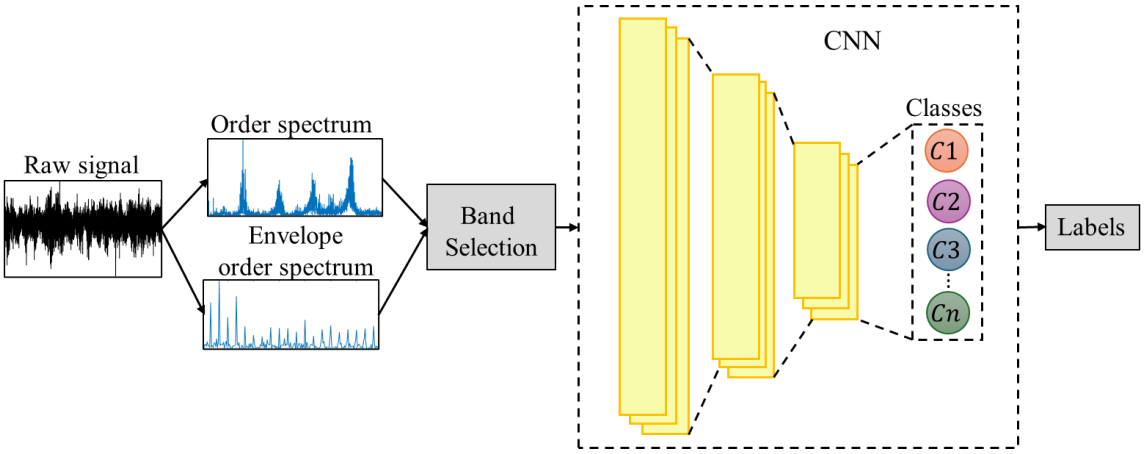


Figure 7.3: Convolutional neural network model using bands from the frequency and envelope spectra in the order domain.

Results for fault diagnosis are presented using the confusion matrix, which reveals the identification accuracy of all categories. The x-axis denotes the predicted fault class, while the y-axis describes the target fault type. The diagonal elements of the confusion matrix represent the number of instances for which the true class matches the predicted class, thus indicating the classification accuracy for each class. Off-diagonal elements, on the other hand, represent misclassifications. Figure 7.4 presents a confusion matrix for a binary classification problem.

In addition, the overall accuracy of the model can be calculated as the sum of the corrected classified samples divided by the total number of samples.

Table 7.1: Architecture of the employed CNNs.

Layer type	Output shape	Number of trainable parameters
Convolutional	(745,12)	156
Max Pooling	(372,12)	0
Convolutional	(181, 24)	3480
Max Pooling	(90, 24)	0
Convolutional	(40, 48)	13872
Max Pooling	(20, 48)	0
Flatten	960	0
Dense	80	76880
Dropout	80	0
SoftMax	6	486

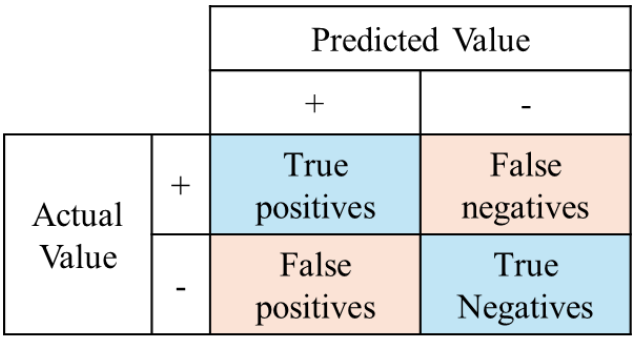


Figure 7.4: Generic confusion matrix of a binary classification problem.

7.2 Results

The first test refers to the capacity of the models to correctly predict faults measured in close conditions to the one used in the training. For this purpose, the models were trained using data corresponding to the working speed of $45\,Hz$ and tested using data measured at 40 and $50\,Hz$. Figures 7.5, 7.6 and 7.7 provide the confusion matrixes for the convolutional neural network using the order spectrum, the convolutional neural network using the frequency bands and the proposed model using domain adaptation, respectively.

The proposed model using domain adaptation achieved an accuracy of 98.42 % demonstrating high capacity in handling domain shifts. On the other hand, the regular convolutional neural network using the order spectrum achieved 72.98% of accuracy ratifying the limited generalization of deep models in the presence of domain variations. Furthermore, the use of bands improved the accuracy of the CNN achieving 83.62%. Additionally, the misclassified data occurs in only one class for the proposed model while there are wrong classifications in several fault types for other two models.

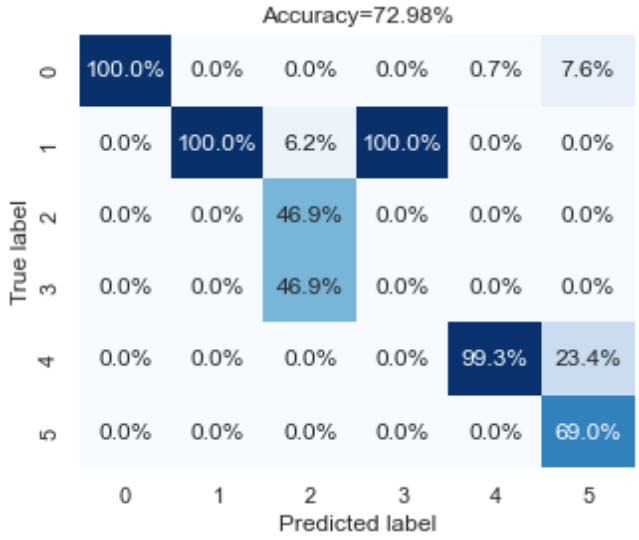


Figure 7.5: Confusion matrix provided by the CNN using the order spectrum for test 1.

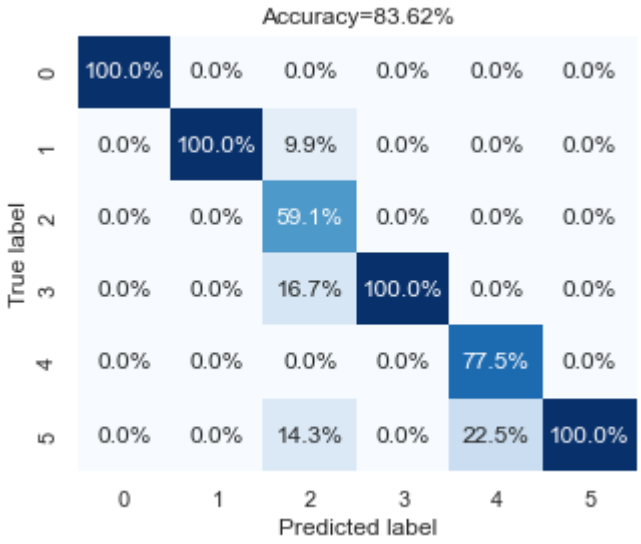


Figure 7.6: Confusion matrix provided by the CNN using frequency bands for test 1.

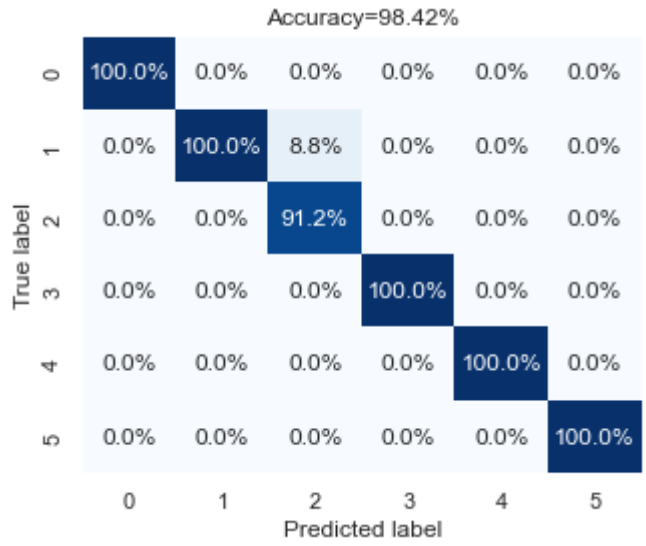


Figure 7.7: Confusion matrix provided by the proposed domain adaptation framework for test 1.

A second test was carried out considering a major variability of the working conditions. The models were trained with data measured at 35 and 45 Hz and tested with samples from $30, 40,$ and 50 Hz. Results are shown in Figures 7.8, 7.9 and 7.10. The accuracy of the three models decreased in comparison to the first test. This is because the data distributions present major differences due to the dynamic response of the machine. For the employed dataset, peaks at fault characteristic frequencies on vibration signatures at 30 and 35 Hz are less defined in comparison to those at 45 and 50 Hz. therefore, the mapping problem is more complex due to these variations.

The convolutional neural network model using the order spectrum achieved an accuracy of 56.19 % exhibiting poor generalization capacity for test data from different speeds to the one used to train. The use of bands again improved the accuracy to 72.54% suggesting that deep learning models present better generalization when prior knowledge about the faults is applied. This can be explained because all information contained in the input vector is meaningful for the fault classification. When the whole spectrum is used, the network must identify what information is important to define every class and which points of the spectrum do not provide relevant characteristics to differentiate each fault pattern.

The proposed domain adaptation model achieved an accuracy of 96.32% and it was the model that presented the lower performance decrease. This can be explained due to how the classifier of the neural networks assigns labels to data. When a neural network is trained in a

supervised manner, the extracted features are rigid. This means that regardless of speed conditions, every class has the same well-defined pattern. Therefore, when new data is fed to the model, the network will try to fit to one of those patterns without considering the possible variations. In contrast, autoencoders extracts different feature vectors for every class in each speed conditions. This enables the comparison of features with the closer speed condition taking advantage of the correlation analysis and then providing better generalization results.

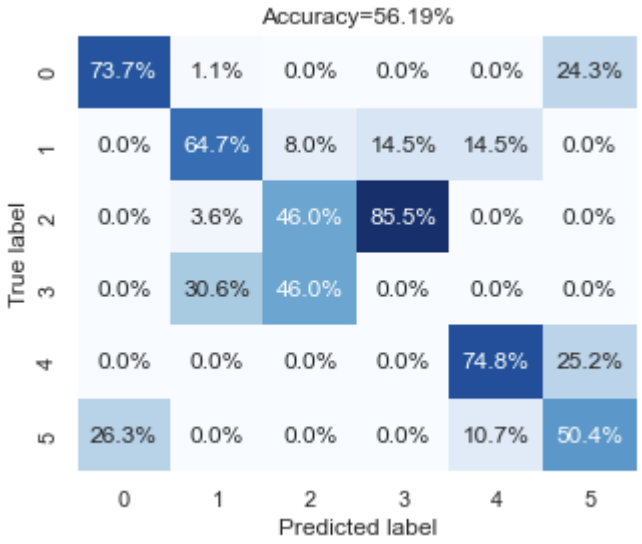


Figure 7.8: Confusion matrix provided by the CNN using the order spectrum for test 2.

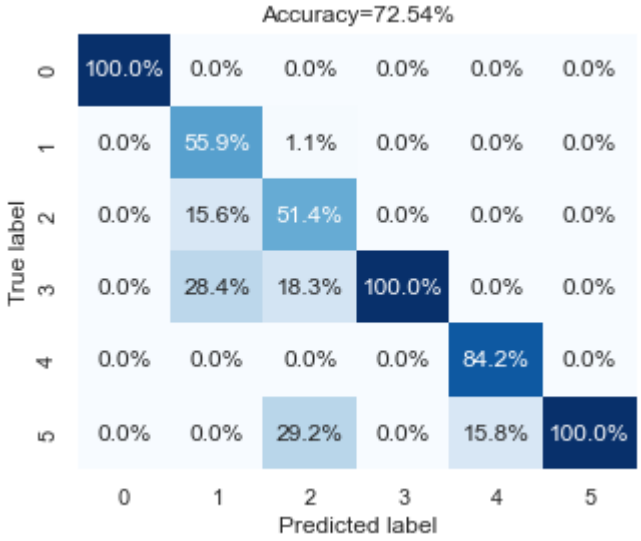


Figure 7.9: Confusion matrix provided by the CNN using frequency bands for test 2.

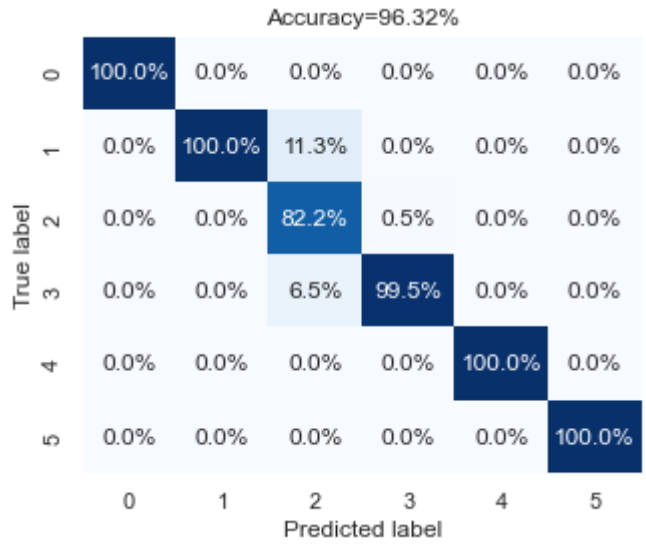


Figure 7.10: Confusion matrix provided by the proposed domain adaptation framework for test 2.

It is important to note that the proposed domain adaptation model and both convolutional neural network models provide outstanding results using testing data belonging to the same speed conditions used to train. Table 7.2 summarizes the accuracies of the models in the two testing scenarios, including outcomes using testing data with the same distribution as those used to train the models.

Table 7.2: Accuracies of the models in two different testing scenarios.

		CNN with spectrum	CNN with frequency bands	Proposed domain adaptation model
Same speed conditions to those of training	Test 1	99.64%	99.95%	99.93%
	Test 2	99.23%	99.71%	99.66%
Different speed conditions to those of training	Test 1	72.98%	83.62%	98.42%
	Test 2	56.19%	72.54%	96.32%

The previous results demonstrate that deep learning models diagnose very well when working conditions do not vary. However, Issues appear in the presence of speed shifts. Data distribution changes due to the dynamic response of the gearbox. Conventional deep learning approaches has limitations to handle with data that present different patterns to those learned in

the training. Therefore, domain adaptation strategies are required to enhance the generalization of the models.

The proposed methodology provided promising results for the gearbox fault diagnosis under variable speed conditions. When compared to existing approaches, the proposed framework overcomes the limitations of label requirements to perform domain adaptation since it is based on an unsupervised reconstruction model. In addition, as the feature synchronization is done by calculating the correlation between the extracted features, the domain adaptation is not limited to only one target domain. Finally, the required computational power is low facilitating its application in real industrial scenarios.

8 CONCLUSIONS

This thesis introduced a novel domain adaptation methodology for gearbox fault diagnosis under variable speed conditions. The framework combined prior knowledge of the fault signatures and deep learning models to extract discriminative features for classification. Domain adaptation at different speed conditions is accomplished by applying the correlation analysis. The methodology was evaluated using experimental data from the PHM 2009 competition. The introduced approach can be used for two applications: A clustering-based diagnosis and a reconstruction-based diagnosis system.

The proposed domain adaptation can be applied for clustering. Experimental results demonstrate the effectiveness of the proposed framework for unsupervised clustering and domain adaptation. The methodology achieved outstanding results in the correct identification of clusters belonging to the same health condition across different working speeds. The convolutional autoencoder extracted features with high correlation enabling the identification of similar patterns in the presence of domain shifts. The global normalization played a significant role in the domain adaptation, avoiding eliminating distinctive characteristics from the input vector essential for the diagnosis of the different fault types. The size of the extracted features is also relevant since the consideration of an insufficient number of dimensions can lead to misclassification due to loss of information. The proposed domain adaptation methodology enables the identification of fault patterns across different speed conditions, overcoming limitations of some existing methods such as the requirement of label from the source domain and the number of classes in each working condition.

Furthermore, the proposed domain adaptation can be applied to the construction of a diagnosis system based on the reconstruction of the input information. Experimental results exhibit the high performance of the proposed framework to diagnose faults under variable working conditions. Tests using data from different speeds to those used to train demonstrate that the introduced method is suitable for domain adaptation. The global accuracy of the proposed model remained excellent even in scenarios considering large variations of data distributions. Comparisons with traditional convolutional neural network models corroborate the need for domain adaptation strategies to improve the generalization and performance of deep learning models.

Finally, the implementation of the methodology is straightforward and the computational power requirements are low in comparison to existing frameworks. The promising results encourage more research to enhance the robustness of the method.

8.1 Suggestion for future work

Considering the promising results of this study, some potential research is suggested to increase the scope of the methodology:

- Evaluate the proposed methodology considering more complex gearbox configurations including more reductions and planetary arrangements.
- Consider data distribution changes due to load fluctuations.
- Study new normalization methodologies to improve the diagnosis of intricate dynamic systems.
- Employ datasets measured in industrial environments containing noise to improve the robustness of the methodology.
- Explore different architectures of models including physics informed neural networks and capsule neural networks.
- Combine the framework with decision-making algorithms to introduce new prior information and enrich the diagnosis.

These suggestions aim to improve the fault diagnosis of gearboxes under variable working conditions and seek application in real industrial scenarios.

REFERENCES

BAO, H.; YAN, Z.; JI, S.; WANG, J.; JIA, S.; ZHANG, G.; HAN, B. An enhanced sparse filtering method for transfer fault diagnosis using maximum classifier discrepancy. **Measurement Science and Technology**, [*s. l.*], v. 32, n. 8, 2021.

CAO, X.; CHEN, B.; ZENG, N. A deep domain adaption model with multi-task networks for planetary gearbox fault diagnosis. **Neurocomputing**, [*s. l.*], v. 409, 2020.

CHEN, Z.; HE, G.; LI, J.; LIAO, Y.; GRYLLIAS, K.; LI, W. Domain Adversarial Transfer Network for Cross-Domain Fault Diagnosis of Rotary Machinery. **IEEE Transactions on Instrumentation and Measurement**, [*s. l.*], v. 69, n. 11, 2020.

CHEN, X.; SHAO, H.; XIAO, Y.; YAN, S.; CAI, B.; LIU, B. Collaborative fault diagnosis of rotating machinery via dual adversarial guided unsupervised multi-domain adaptation network. **Mechanical Systems and Signal Processing**, [*s. l.*], v. 198, 2023.

CHEN, C.; SHEN, F.; XU, J.; YAN, R. Domain Adaptation-Based Transfer Learning for Gear Fault Diagnosis under Varying Working Conditions. **IEEE Transactions on Instrumentation and Measurement**, [*s. l.*], v. 70, 2021.

CHEN, J.; WANG, J.; ZHU, J.; LEE, T. H.; DE SILVA, C. W. Unsupervised Cross-Domain Fault Diagnosis Using Feature Representation Alignment Networks for Rotating Machinery. **IEEE/ASME Transactions on Mechatronics**, [*s. l.*], v. 26, n. 5, 2021.

DA SILVA, I. N.; SPATTI, D. H.; FLAUZINO, R. A.; LIBONI, L. H. B.; DOS REIS ALVES, S. F. **Artificial neural networks: A practical course**. [*S. l.: s. n.*], 2016. 2016.

FAN, Z.; XU, Q.; JIANG, C.; DING, S. X. Weighted quantile discrepancy-based deep domain adaptation network for intelligent fault diagnosis. **Knowledge-Based Systems**, [*s. l.*], v. 240, 2022.

GHIFARY, M.; BASTIAAN KLEIJN, W.; ZHANG, M. Domain adaptive neural networks for object recognition. **Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)**, [*s. l.*], v. 8862, 2014.

HE, Z.; SHAO, H.; WANG, P.; LIN, J. (Jing); CHENG, J.; YANG, Y. Deep transfer multi-wavelet auto-encoder for intelligent fault diagnosis of gearbox with few target training samples. **Knowledge-Based Systems**, [*s. l.*], v. 191, 2020. Disponível em: https://doi.org/10.1016/j.knosys.2019.105313.

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**, [*s. l.*], v. 7, p. 115368–115377, 2019.

HUANG, D.; ZHANG, W. A.; GUO, F.; LIU, W.; SHI, X. Wavelet Packet Decomposition-Based Multiscale CNN for Fault Diagnosis of Wind Turbine Gearbox. **IEEE Transactions on Cybernetics**, [*s. l.*], v. 53, n. 1, 2023.

JAMES I. TAYLOR. **The Vibration Analysis Handbook: A Practical Guide for Solving Rotating Machinery Problems**. 2. ed. [*S. l.: s. n.*], 2003. 2003. Disponível em: https://books.google.com.br/books/about/The_Vibration_Analysis_Handbook.html?id=LvTOnQEACAAJ&redir_esc=y. Acesso em: 4 jun. 2024.

JANG, G. B.; CHO, S. B. Cross-Domain Adaptation Using Domain Interpolation for Rotating Machinery Fault Diagnosis. **IEEE Transactions on Instrumentation and Measurement**, [*s. l.*], v. 71, 2022.

JIA, M.; WANG, J.; ZHANG, Z.; HAN, B.; SHI, Z.; GUO, L.; ZHAO, W. A novel method for diagnosing bearing transfer faults based on a maximum mean discrepancies guided domain-adversarial mechanism. **Measurement Science and Technology**, [*s. l.*], v. 33, n. 1, 2022.

JIANG, X.; WANG, X.; HAN, B.; WANG, J.; ZHANG, Z.; MA, H.; XING, S.; MAN, K. A novel hybrid distance guided domain adversarial method for cross domain fault diagnosis of gearbox. **Measurement Science and Technology**, [*s. l.*], v. 34, n. 6, 2023.

JIAO, J.; LIN, J.; ZHAO, M.; LIANG, K. Double-level adversarial domain adaptation network for intelligent fault diagnosis. **Knowledge-Based Systems**, [*s. l.*], v. 205, 2020a.

JIAO, J.; ZHAO, M.; LIN, J.; LIANG, K. Residual joint adaptation adversarial network for intelligent transfer fault diagnosis. **Mechanical Systems and Signal Processing**, [*s. l.*], v. 145, p. 106962, 2020b. Disponível em: https://doi.org/10.1016/j.ymssp.2020.106962.

KIM, M.; KO, J. U.; LEE, J.; YOUN, B. D.; JUNG, J. H.; SUN, K. H. A Domain Adaptation with Semantic Clustering (DASC) method for fault diagnosis of rotating machinery. **ISA Transactions**, [*s. l.*], v. 120, 2022.

KIM, T.; LEE, S. A Novel Unsupervised Clustering and Domain Adaptation Framework for Rotating Machinery Fault Diagnosis. **IEEE Transactions on Industrial Informatics**, [*s. l.*], v. 19, n. 9, 2023.

LECUN, Y.; BOTTOU, L.; BENGIO, Y.; HAFFNER, P. Gradient-based learning applied to document recognition. **Proceedings of the IEEE**, [*s. l.*], v. 86, n. 11, p. 2278–2323, 1998.

LEE, S.; KIM, S.; KIM, S. J.; LEE, J.; YOON, H.; YOUN, B. D. Revolution and peak discrepancy-based domain alignment method for bearing fault diagnosis under very low-speed conditions. **Expert Systems with Applications**, [*s. l.*], v. 251, p. 124084, 2024. Disponível em: Acesso em: 20 maio 2024.

LEE, J.; KO, J. U.; KIM, T.; KIM, Y. C.; JUNG, J. H.; YOUN, B. D. Domain adaptation with label-aligned sampling (DALAS) for cross-domain fault diagnosis of rotating machinery under class imbalance. **Expert Systems with Applications**, [*s. l.*], v. 243, 2024.

LI, X.; JIA, X. D.; ZHANG, W.; MA, H.; LUO, Z.; LI, Xu. Intelligent cross-machine fault diagnosis approach with deep auto-encoder and domain adaptation. **Neurocomputing**, [*s. l.*], v. 383, 2020.

LI, R.; LI, S.; XU, K.; LI, X.; LU, J.; ZENG, M.; LI, M.; DU, J. Adversarial domain adaptation of asymmetric mapping with CORAL alignment for intelligent fault diagnosis. **Measurement Science and Technology**, [*s. l.*], v. 33, n. 5, 2022.

LI, Y.; REN, Y.; ZHENG, H.; DENG, Z.; WANG, S. A Novel Cross-Domain Intelligent Fault Diagnosis Method Based on Entropy Features and Transfer Learning. **IEEE Transactions on Instrumentation and Measurement**, [*s. l.*], v. 70, 2021.

LI, Q.; TANG, B.; DENG, L.; WU, Y.; WANG, Y. Deep balanced domain adaptation neural networks for fault diagnosis of planetary gearboxes with limited labeled data. **Measurement: Journal of the International Measurement Confederation**, [*s. l.*], v. 156, p. 107570, 2020. Disponível em: https://doi.org/10.1016/j.measurement.2020.107570.

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**, [*s. l.*], v. 159, p. 111620, 2024. Disponível em: Acesso em: 20 jun. 2024.

LIU, Z. H.; JIANG, L. B.; WEI, H. L.; CHEN, L.; LI, X. H. Optimal Transport-Based Deep Domain Adaptation Approach for Fault Diagnosis of Rotating Machine. **IEEE Transactions on Instrumentation and Measurement**, [*s. l.*], v. 70, 2021.

LIU, S.; JIANG, H.; WU, Z.; YI, Z.; WANG, R. Intelligent fault diagnosis of rotating machinery using a multi-source domain adaptation network with adversarial discrepancy matching. **Reliability Engineering and System Safety**, [*s. l.*], v. 231, 2023.

LOU, Y.; KUMAR, A.; XIANG, J. Machinery Fault Diagnosis Based on Domain Adaptation to Bridge the Gap Between Simulation and Measured Signals. **IEEE Transactions on Instrumentation and Measurement**, [*s. l.*], v. 71, 2022.

LU, W.; LIANG, B.; CHENG, Y.; MENG, D.; YANG, J.; ZHANG, T. Deep Model Based Domain Adaptation for Fault Diagnosis. **IEEE Transactions on Industrial Electronics**, [*s. l.*], v. 64, n. 3, 2017.

MCCULLOCH, W. S.; PITTS, W. A logical calculus of the ideas immanent in nervous activity. **The Bulletin of Mathematical Biophysics**, [*s. l.*], v. 5, n. 4, 1943.

NI, Q.; JI, J. C.; HALKON, B.; FENG, K.; NANDI, A. K. Physics-Informed Residual Network (PIResNet) for rolling element bearing fault diagnostics. **Mechanical Systems and Signal Processing**, [*s. l.*], v. 200, p. 110544, 2023. Disponível em: Acesso em: 22 abr. 2024.

PANG, S. Stacked maximum independence autoencoders: A domain generalization approach for fault diagnosis under various working conditions. **Mechanical Systems and Signal Processing**, [*s. l.*], v. 208, 2024.

PANG, X.; XUE, X.; JIANG, W.; LU, K. An Investigation into Fault Diagnosis of Planetary Gearboxes Using A Bispectrum Convolutional Neural Network. **IEEE/ASME Transactions on Mechatronics**, [*s. l.*], v. 26, n. 4, 2021.

PROGNOSTICS AND HEALTH MANAGEMENT SOCIETY. **PHM 2009 competition dataset**. [*S. l.: s. n.*], [*s. d.*].

QIAN, Q.; LUO, J.; QIN, Y. Adaptive Intermediate Class-Wise Distribution Alignment: A Universal Domain Adaptation and Generalization Method for Machine Fault Diagnosis. **IEEE Transactions on Neural Networks and Learning Systems**, [*s. l.*], 2024.

QIAN, Q.; WANG, Y.; ZHANG, T.; QIN, Y. Maximum mean square discrepancy: A new discrepancy representation metric for mechanical fault transfer diagnosis. **Knowledge-Based Systems**, [*s. l.*], v. 276, 2023.

QIN, Y.; QIAN, Q.; LUO, J.; PU, H. Deep Joint Distribution Alignment: A Novel Enhanced-Domain Adaptation Mechanism for Fault Transfer Diagnosis. **IEEE Transactions on Cybernetics**, [*s. l.*], v. 53, n. 5, 2023a.

QIN, Y.; QIAN, Q.; WANG, Z.; MAO, Y. Adaptive manifold partial domain adaptation for fault transfer diagnosis of rotating machinery. **Engineering Applications of Artificial Intelligence**, [*s. l.*], v. 126, 2023b.

QIN, Y.; YAO, Q.; WANG, Y.; MAO, Y. Parameter sharing adversarial domain adaptation networks for fault transfer diagnosis of planetary gearboxes. **Mechanical Systems and Signal Processing**, [*s. l.*], v. 160, 2021.

RAN, M.; TANG, B.; SUN, P.; LI, Q.; SHI, T. A gradient aligned domain adversarial network for unsupervised intelligent fault diagnosis of gearboxes. **ISA Transactions**, [*s. l.*], v. 148, p. 461–476, 2024. Disponível em: Acesso em: 21 maio 2024.

REZAEIANJOUYBARI, B.; SHANG, Y. A novel deep multi-source domain adaptation framework for bearing fault diagnosis based on feature-level and task-specific distribution alignment. **Measurement: Journal of the International Measurement Confederation**, [*s. l.*], v. 178, 2021.

ROSENBLATT, F. The perceptron: A probabilistic model for information storage and organization in the brain. **Psychological Review**, [*s. l.*], v. 65, n. 6, 1958.

SCHWENDEMANN, S.; AMJAD, Z.; SIKORA, A. Bearing fault diagnosis with intermediate domain based Layered Maximum Mean Discrepancy: A new transfer learning approach. **Engineering Applications of Artificial Intelligence**, [*s. l.*], v. 105, 2021.

SHI, Y.; DENG, A.; DING, X.; ZHANG, S.; XU, S.; LI, J. Multisource domain factorization network for cross-domain fault diagnosis of rotating machinery: An unsupervised multisource domain adaptation method. **Mechanical Systems and Signal Processing**, [*s. l.*], v. 164, 2022.

SU, Z.; ZHANG, J.; TANG, J.; WANG, Y.; XU, H.; ZOU, J.; FAN, S. A novel deep transfer learning method with inter-domain decision discrepancy minimization for intelligent fault diagnosis. **Knowledge-Based Systems**, [*s. l.*], v. 259, 2023.

TIAN, J.; HAN, D.; KARIMI, H. R.; ZHANG, Y.; SHI, P. Deep learning-based open set multi-source domain adaptation with complementary transferability metric for mechanical fault diagnosis. **Neural Networks**, [*s. l.*], v. 162, 2023.

WANG, R.; HUANG, W.; WANG, J.; SHEN, C.; ZHU, Z. Multisource Domain Feature Adaptation Network for Bearing Fault Diagnosis under Time-Varying Working Conditions. **IEEE Transactions on Instrumentation and Measurement**, [*s. l.*], v. 71, 2022.

WANG, X.; SHE, B.; SHI, Z.; SUN, S.; QIN, F. Partial adversarial domain adaptation by dual-domain alignment for fault diagnosis of rotating machines. **ISA Transactions**, [*s. l.*], 2022.

WANG, R.; YAN, F.; YU, L.; SHEN, C.; HU, X. Joint Wasserstein distance matching under conditional probability distribution for cross-domain fault diagnosis of rotating machinery. **Mechanical Systems and Signal Processing**, [*s. l.*], v. 210, 2024.

WEN, H.; GUO, W.; LI, X. A novel deep clustering network using multi-representation autoencoder and adversarial learning for large cross-domain fault diagnosis of rolling bearings. **Expert Systems with Applications**, [*s. l.*], v. 225, 2023.

WU, F.; LEE, J. Information reconstruction method for improved clustering and diagnosis of generic gearbox signals. **International Journal of Prognostics and Health Management**, [*s. l.*], v. 2, n. 1, 2011.

XIAO, D.; QIN, C.; YU, H.; HUANG, Y.; LIU, C.; ZHANG, J. Unsupervised machine fault diagnosis for noisy domain adaptation using marginal denoising autoencoder based on acoustic

signals. **Measurement: Journal of the International Measurement Confederation**, [*s. l.*], v. 176, 2021.

XIONG, P.; TANG, B.; DENG, L.; ZHAO, M.; YU, X. Multi-block domain adaptation with central moment discrepancy for fault diagnosis. **Measurement: Journal of the International Measurement Confederation**, [*s. l.*], v. 169, 2021.

YANG, S.; KONG, X.; WANG, Q.; LI, Z.; CHENG, H.; XU, K. Deep multiple auto-encoder with attention mechanism network: A dynamic domain adaptation method for rotary machine fault diagnosis under different working conditions. **Knowledge-Based Systems**, [*s. l.*], v. 249, 2022.

YU, Y.; ZHAO, J.; TANG, T.; WANG, J.; CHEN, M.; WU, J.; WANG, L. Wasserstein distance-based asymmetric adversarial domain adaptation in intelligent bearing fault diagnosis. **Measurement Science and Technology**, [*s. l.*], v. 32, n. 11, 2021.

YUAN, S. Z.; LIU, Z. H.; WEI, H. L.; CHEN, L.; LV, M. Y.; LI, X. H. A Variational Auto-Encoder-Based Multisource Deep Domain Adaptation Model Using Optimal Transport for Cross-Machine Fault Diagnosis of Rotating Machinery. **IEEE Transactions on Instrumentation and Measurement**, [*s. l.*], v. 73, 2024.

ZHANG, Z.; CHEN, H.; LI, S.; AN, Z. Sparse filtering based domain adaptation for mechanical fault diagnosis. **Neurocomputing**, [*s. l.*], v. 393, 2020.

ZHAO, K.; JIA, F.; SHAO, H. A novel conditional weighting transfer Wasserstein auto-encoder for rolling bearing fault diagnosis with multi-source domains. **Knowledge-Based Systems**, [*s. l.*], v. 262, 2023.

ZHAO, B.; ZHANG, X.; ZHAN, Z.; WU, Q. Deep multi-scale adversarial network with attention: A novel domain adaptation method for intelligent fault diagnosis. **Journal of Manufacturing Systems**, [*s. l.*], v. 59, 2021.

ZHENG, H.; YANG, Y.; YIN, J.; LI, Y.; WANG, R.; XU, M. Deep Domain Generalization Combining A Priori Diagnosis Knowledge Toward Cross-Domain Fault Diagnosis of Rolling Bearing. **IEEE Transactions on Instrumentation and Measurement**, [*s. l.*], v. 70, p. 1–11, 2021.