Chapter 1

Introduction

Industry 4.0 has arisen as a promising technology framework that introduced increased automation in the manufacturing process to improve efficiency, reliability, and availability. Consequently, the production process has been integrated and extended at both intra- and inter-organizational levels, increasing the complexity [1–3]. This increased the cost and challenge of ensuring the integrity of machinery by continuous health monitoring and performance degradation analysis. The rise of machine learning profoundly impacted the production process in Industry 4.0 [4–7]. Meantime, RM is the heart of modern mechanical and electro-mechanical systems as it roughly makes up 40.0% of all manufacturing machinery. However, due to moving or rotating components, it and is more likely to go through both deterioration and failures [8]. Hence the industrial applications are primarily concerned about the sustainability and performance of RM to a great extent. This necessitates employing periodic maintenance in a predictive or preventive manner without affecting the regular operation of the machines. Equipment deterioration mainly happens in RM by aging or improper maintenance, which increases the risk of failure and affects the system's efficient functioning. Suppose proper fault diagnosis and predictive maintenance strategy have not been developed for a system. In that case, the unexpected failures affect productivity, availability, reliability, safety, etc., inviting new researchers to establish intelligent methods and tools. Condition-based maintenance plays a critical role in the maintenance strategy building process and is mainly performed by vibration signal monitoring and analysis. The heterogeneous sensors deployed in the machinery provide a sufficient amount of real-time data to analyze the system's present condition. The continuous monitoring of such data helps predict the fault at an early stage such that preventive measures can be taken before the whole system shuts down unexpectedly.

Various physical magnitudes such as stator current, vibration, and sound signals can be used to base the condition monitoring applied to RM. That said, the most accepted strategies are vibration-based monitoring and motor current signature analysis, where vibration analysis has been used for completing above 82.0% of fault diagnosis methodologies [3]. The dynamic forces within an RM produce a vibration force, and even at an early stage, the vibration pattern changes when a fault is observed. Due to this property, RM fault diagnosis with vibration analysis became very popular in the last decade. It can therefore be said that vibration analysis is an efficient and trustworthy tool for assessing the machinery condition [9–11].

The RM can be primarily segregated into three parts at the component level: bearings, gears, and rotors [12]. Even though there is an abundance of written work related to bearing and gear fault identification, the faults that affect the rotor system are addressed by relatively fewer works [13–15]. In this respect, the literature related to rotor fault diagnosis (RFD) looks a little fragmental. It thus falls short of providing opportunities for using the fault-specific characteristics of RM faults and making significant research improvements within the domain using the feature engineering phase of AI. Structural rotor faults (SRF) are the rotor faults that have a direct, even catastrophic impact on the structural attributes and performance of the equipment that has been affected. Not only that, surrounding equipment like the bearings and gears may also be faced with secondary faults [16] as a result of SRF. This makes the analysis of rotor faults preponderant. This thesis, therefore, attempts to draw more AI research attention towards RFD, especially the attendant properties and SRF. There are overlooked loopholes in the utilization of the fault characteristics and the associated features as most of the studies related to RM fault diagnosis give emphasis to component-wise analysis instead of focusing on fault-wise analysis [17].

In this regard, this study aims at finding a suitable fault-wise categorization of rotor faults based on their vibrational characteristics as a primary step. Such a characterization is important because the previous studies on rotor faults make scarce reference to the fault characteristics, the fault correlation, fault simulation process, etc. Similarly, an ideal AI-based RFD framework comprising the data acquisition, feature processing, and classification steps in ML and DL perspectives is proposed in the study. A general summary of the state-of-the-art techniques, methods, and algorithms used in different phases is provided in the framework. Along with this, the significant rotor faults, their vibrational frequency and phase characteristics, the causes and effects, the associated faults, and practical aspects are also analyzed in the thesis. Rotor faults, SRFs in particular, are more sensitive to DFC. Hence, in the current scenario, it is highly recommended to emphasize research that uses these fault-specific DFC with AI. The thesis focused on the research advancements, including the impact of DFC in various phases of the framework. In this direction, the significance of DFC in SRF decision making is showed with various ML techniques, including fuzzy logic.

Due to inadequate and unrealistic faulty data, most of the AI-based solutions in RFD are presently limited to an experimental stage. The scope of employing advanced DL strategies is, therefore, similarly limited. It becomes difficult to produce generalized solutions for RFD as the data gathered from the real-world industrial scenario is extremely imbalanced. Data augmentation is a solution to this issue, but the commonly used augmentation methods do not consider the TS characteristics of the signal which generates unrealistic synthesized data. We, therefore, create an SRF diagnosis framework, which develops a subsampled data set incorporating DFC to address the issues of industrial data acquisition. An augmentation method using dynamic time warping (DTW) that is updated by fault information content-based weighing scheme is used for handling the problems related to data scarcity and imbalance. For the improvement of the scope of the projected SRF diagnosis solutions, the augmentation provided more heterogeneity and discriminative features to the synthesized samples. The faults are predicted together with a standard tradeoff between earliness and accuracy through the early classification approach for SRF. Towards this end, a sequential deep learning classifier is first of all created by considering accuracy only as an objective before moving on to early decision policy that is defined by taking accuracy and earliness into account. Popular sequential learning methods such as long short-term memory (LSTM) and gated recurrent unit (GRU) have been utilized to make the decision with the partially observed data. Thus the early class prediction showed its significance in fault diagnosis of SRF by making fact-based decisions at the earliest possible without having to wait for full-length.

Similarly, even though various imaging schemes are present in the literature regarding the 2D representation of vibration data to make it compatible with convolutional neural networks (CNN), most of them do not consider the TS characterization of the input data. This generates extraneous data representation for further feature processing or decision making that often ends up with far-fetched results. In such a scenario, this study experimented with popular TS imaging techniques like recurrence plot (RP), Gramian angular field (GAF), and Markov transition field (MTF). Finally, RP has been selected for further fault processing due to its exceptional fault diagnosis ability with SRF data compared to the other two methods. RP is set as the primary imaging scheme, and it is modified with fuzzy logic to generate smoother fuzzy-RP (FRP) to produce more accurate results. Then FRP was presented with two variants combining individual sensor FRPs based on its fault pattern matching score assisted by DFC.

Such a scheme is compared with basic RP and other RFD techniques to prove its significance. Finally, advanced DL methods have been used to compare the different SRF data representations and diagnosis strategies. Similarly, it is derived from the writings that some drawbacks are reinforced by fault diagnosis depended on a single information source and a distinct approach. Hence we proposed a classifier fusion scheme assisted with fuzzy integral (FI), that depends on the confidence of the sources generated by parallel decision-making strategies such as CNN and LSTM. The suggested fusion of the decision scores based on CNN together with the sequential classifier like LSTM based decision scores hugely improved the outcomes across the board.

Lastly, we utilized the attention mechanism with transformer architecture to identify both long-term and short-term dependencies of the sensor signals in complex RM. With that, the attention score assisted with fault pattern-based ranking used for sensor fusion to ensure the relative importance of fused sensor vectors and their fault sensitivity. The experiments showed the transformer's ability to adapt to the fault diagnosis domain and deal with complex rotor systems with multiple sensors. Besides the generic transformer, the recurrent transformer models are also applied to assess the local dependency among the embedding vectors. The proposed method is capable of capturing sufficient dependency information even from short-length sequences, with fewer dimension embedding, thereby significantly reducing the execution time. The experimental results signify the effectiveness of the proposed framework in utilizing transformers with multi-sensor data, incorporating fault information content, and excelling in various industrial working environments.

All the developed methods are tested on two SRF data sets. The primary data set is a novel SRF data set that has been developed by a testbed and simulation of industrial scenarios. To add to that, a data set with varying speed and load conditions that is available publicly has also been used, and the model's performance is compared with advanced learning methods. The overall results from the experiments show that all the proposed methods enhance the accuracy of fault classification and promote more research in SRF diagnosis.

1.1 Problem Formulation

Developing a reliable machine health monitoring system requires responsiveness towards varying operating conditions and must be robust. It demands data pattern analysis to utilize the most sensitive symptomatic fault features, proper data representation to enhance input data/features characteristics, and advanced decision-making strategies. This poses several challenges for researchers as well as practitioners. Thus, this work aims to identify and utilize the DFC of vibration pattern in various input data representations and classification models while employing the most advanced learning strategies addressing the challenges of the practical working environment.

1.2 Scope of the Work

The objective of FDPM is to predict the fault in advance by continuously monitoring the system's condition and executing the maintenance policy. The observation, decision-making on faults, and periodic maintenance require sufficient prior knowledge about the system and technical expertise. It includes personal subjectivity and other implementational challenges, which increase production cost and time and the possibility of false defect detection. Moreover, the fault detection based on threshold values lacks the opportunity of learning from experience and often fails in unforeseen situations and turns out to be an unprecise approach. In such a scenario, a robust FDPM system with domain expertise, fast processing, and self-learning capability is essentially the need of the hour in industrial production. Intelligent fault diagnosis methods and tools must be developed to satisfy the diagnosis requirement based on data collected for conditional monitoring. Nowadays, vibration signal monitoring has been extensively employed in real-time condition monitoring of modern systems to protect from unexpected shutdowns and catastrophic failures without interrupting the machine operations at an incipient stage. Furthermore, with the relatively harsh working conditions and continuous operations, the rotating components of machinery are more prone to early-stage defects that lead to the reduced lifetime of the system. But in the current scenario, the RM fault diagnosis has been revolved around gear or bearing faults overlooking the rotor faults, which are the root cause that affects the structural components and triggers secondary faults like bearing and gear faults. However, the research community in RFD precluded the advantages of utilizing DFC, applying the sequential properties of raw vibration signal, and exploring sequential DL models. Similarly, the literature is fragmentary as it lacks advanced learning strategies and fails in benefitting from different input representation methods. The critical challenge of RM fault diagnosis is attracting more research into the domain with a proper fault-wise characterization of the defect, addressing the root cause. Thus this thesis aims to uphold the importance of RFD in RM and develop the diagnosis framework with enhanced accuracy and responsiveness. The developed methods succeeded in addressing the data-related problems of the industrial environment, employing various data representation methods, identifying and utilizing fault-specific features in different phases, and exploring advanced learning strategies. The proposed methods have been tested with two different datasets simulating varying operational conditions of the industrial environment.

1.3 Objectives

- Identify fault specific distinctive frequency components and utilize them in various phases of the framework to develop more domain-specific solutions.
- To address the data acquisition related issues of real plant data and deal with different working conditions using data subsampling and TS data augmentation and sensor fusions.

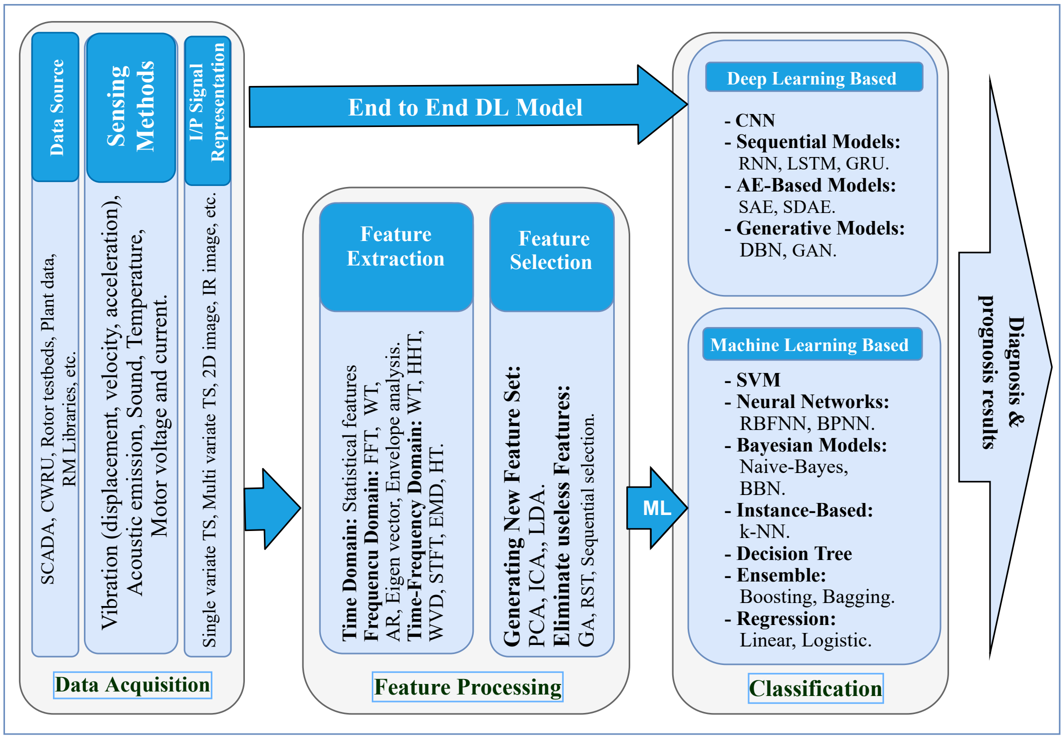

Figure 1.1: AI-based RFD framework

- Identify the most appropriate input data representation for raw vibration, to keep the characteristics of the input and present to decision-making phase.
- Apply the sequential deep learning models and advanced learning strategies to provide enhanced results.

1.4 Methodology

The models proposed in this thesis follow a generalized framework for SRF, demonstrated in Fig. 1.1. Even though several types of inputs are shown in the framework, raw vibration data is typically used as the primary source of information for SRF and has been widely used for other fault diagnoses. A model AI-based framework for RFD is made up of three primary stages: data acquisition, feature processing, and classification. The ML part follows the abovementioned stages in a sequential manner, while DL leaves out the feature processing phase because of its built-in feature learning capability. The dataset is divided into training, validation, and testing data, where the learning process uses training data to learn the parameters of the model. In contrast, validation data gives the estimate of trained parameters while the learning is going on. The final evaluation of the model is performed with the testing data after the training has been completely finished. The detailed description of the three phases of the framework is as follows:

1.4.1 Data acquisition

The data source, acquired signal type, and data representation are defined in the data acquisition phase of RFD. The different methods used in each of these are described below.

Data source: The data collection using rotor testbeds (RTB) is the reasonable and omnipresent method used by the research community for gathering fault data. We used the proposed TS augmentation techniques to handle data imbalance and scarcity issues. Two vibration data sets are used to evaluate the proposed methods.

Sensing method: Vibration has become the most popular signal sensing method for RFD as the nature and behavior of the vibration are affected by the faults related to the rotor. Although vibration signals transport the system's dynamic information, they are sensitive to noise and the sensor mount positions. In scenarios like these, alternate sensing methods are used for data acquisition. For instance, voltage and current (V&C) [18] sensing is more prevalent than vibration sensing in induction motor fault analysis. Many researchers, in the meantime, have given the emphasis on utilizing various other sensing methods such as acoustic emission (AE) [19], sound [20], and temperature [21]. In this relation, combined signals were also utilized, which are created by either simple combination of signals or by the fusion of two or more types of signals [19, 22].

Input representation: The vibration signals are generally represented in a univariate or multivariate TS, and subsampled DFC stream in the experiments. That said, the TS imaging techniques (FRP, GAF, and MTF) are used during their application to CNN model.

1.4.2 Feature processing

Suppression of the noise, identification of the fault-specific features, and presentation of only the essential features to the subsequent phase are the tasks associated with the feature processing phase. In order to extract the information buried in the input data, a range of different signal processing techniques is performed. The overall tasks are performed in two steps: feature extraction and feature selection.

The methods followed for feature extraction are primarily categorized into the time-domain (TD), frequency-domain (FD), and time-frequency-domain (TFD) techniques. Here, the statistical features [23] are commonly operated in the TD feature extraction process. In the FD, fast Fourier transform (FFT) [24], autoregressive (AR) model [25], eigenvector, and envelope analysis are the frequently applied. The methods which include short-time Fourier transform (STFT) [26], empirical mode decomposition (EMD), Hilbert–Huang transform (HHT), Hilbert transform (HT) [27], and wavelet transform (WT) [28] are popularly used in TFD category. The feature selection is done either by generating new features or by eliminating non-relevant features from the existing feature-set. The popular dimensionality reduction techniques, which include principal component analysis (PCA) [29], linear discriminant analysis (LDA) [30], and independent component analysis (ICA) [20], fall within the first category, while rough set theory (RST) [27], genetic algorithm (GA) [27] and sequential selection (SS) [31] are used in the second category methods.

From the framework, it is apparent that ML-based models make use of the features provided by the prior phase. On the other hand, raw data is accepted as input in DL-based models. We used FFT and statistical feature processing for generating DFC and subsampled datasets in the proposed models. The description of DFC feature extraction is provided in section 3.4.

1.4.3 Fault classification

ML- or DL-based models are used for the classification/prediction in the ultimate phase of the AI-based framework. The most widely applied ML classifiers for RFD are the support vector machine (SVM) model [32], the artificial neural network (ANN) model [33], and a number of their variants. Under the instance-based category, the k-Nearest Neighbor (k-NN) [34], and among the probability-based Bayesian methods - naïve-Bayes (NB) [35] and the Bayesian belief network (BBN) [36] – are also demonstrated in relation to RFD. The non-parametric type –which includes decision tree (DT) [37], random forest (RF) algorithms [38] – as well as simple classifiers such as logistic regression (LR) [38] and linear discriminant analysis (LDA), [39] also feature within the literature. A number of researchers have also introduced the AdaBoost (AB) [40] algorithm and other ensemble classification algorithms [41]. The CNNs [42] have attracted more attention than any other DL model in RFD. The autoencoder (AE)-based models, such as stacked AE (SAE) [43] and stacked denoise AE (SDAE) [44], also find their place in the literature. Deep belief networks (DBN) [45] are also widely used in this context. The sequential DL models, such as a recurrent neural network (RNN) and the variants, known as LSTM [46] and GRU [47], have also been explored in relation to RFD. Meanwhile, the deep generative model, such as the generative adversarial network (GAN), is also present in RFD literature.

We used sequential DL models as we prefer maintaining the sequential property of the data in all the phases. Moreover, we incorporated advanced methods such as early classification strategy, decision-level classifier fusion, and attention-based classification in this phase.

1.5 Significant Contributions

The key contributions of this thesis work are summarized as follows:

Identifying the importance of the vibration signals in RM faults, a novel rotor fault dataset of vibration generated artificially in a supervised manner using a rotor testbed. The industrial data acquisition environments are simulated, providing various rotor speeds and defects, including runup and rundown conditions.

This work provided an appropriate categorization of RM faults to separate rotor faults from the bearing and gear faults and categorized them from a 'cause of vibration' perspective, encouraging fault-characteristics-based analysis beyond traditional component-wise analysis.

The research on RM faults is performed such that it addresses the root cause affecting the structural integrity of RM. Hence the structural faults are characterized by their vibrational behavior and grouped and treated under the SRF category.

All the research works identified and emphasized the fault-specific characteristic components such as DFC in various phases of the framework and provided enhanced performance, avoiding the complicated feature generation and selection processes.

The thesis proposed an augmentation method using DTW, enhanced by a fault information content-based weighing scheme to deal with data scarcity and imbalance problems. Data heterogeneity and discriminative ability have been added to data to provide a more generalized SRF solution, and any multi-variate TS vibration dataset can be enhanced with the proposed scheme.

In order to find the most appropriate input image representation for vibration data, the thesis applied three popular TS imaging schemes and identified the predominance of RP. An enhancement on multiple sensors RPs is proposed to generate two categories of FRP images. These imaging schemes are verified by the performance comparison done on the prescribed datasets to show their dominance in SRF.

Two different multi-sensor fusions are proposed in the works based on the fault information content of sensor signals by using a fault pattern matching score. One scheme used for selecting most informative sensor FRPs and other scheme used in embedded sensor segment generation for transformers.

The advanced DL method, such as early classification, is proposed to use for SRF diagnosis to make the decision with the partially observed data with DFC parameters. Popular sequential learning models such as LSTM and GRU have been utilized to produce state-of-the-art results in SRF.

Fuzzy integral based classifier fusion is proposed to combine two parallel decision-making strategies such as CNN and LSTM for SRF diagnosis, and the impact of decision-fusion in the domain has been studied. The performance of classifier fusion with other popular methods has been compared in this work.

Attention mechanism with general transformer architecture and the recurrent transformer model has been utilized to identify both long-term and short-term dependencies of the sensor signals. The experiments showed the transformer's ability to adapt to the fault diagnosis domain and deal with complex rotor systems with multiple sensors.

1.6 Organization of the Thesis

This thesis has been organized into the following seven chapters.

Chapter 1 presents the introduction to rotor fault diagnosis, and structural rotor faults, emphasizing their relevance. It provides an overview of the methodology utilized for fault diagnosis along with the aim and objective of the research work and a summary of significant contributions.

Chapter 2 provides a detailed description of structural rotor faults, including fault categorization, the theoretical background of rotor faults, and existing fault diagnosis approaches with a literature survey of popular ML and DL methods in RFD. This chapter summarized these approaches and presented the research gaps in SRF diagnosis.

Chapter 3 explored the SRF vibration data acquisition process with testbed implementation of rotor faults and provided a brief description of the datasets used in this research. This chapter also describes the relevance of DFC in SRF, the DFC extraction process, and the experimental result of SRF diagnosis with DFC.

Chapter 4 introduces the TS augmentation scheme for vibration data based on soft-DTW. Then the performance of the early classification model is analyzed in detecting SRF, based on a partial observed sequence in a real-time environment.

Chapter 5 describes the use of RP and FRP as a unique image data representation method with TS characterization and system dynamics consideration for multi-sensor vibration data. Also, it examined the impact of classifier fusion of CNN and LSTM with fuzzy integral fusion.

Chapter 6 showed the applicability of the attention mechanism with transformers in SRF diagnosis by proposing a domain-specific embedding representation. It introduced multi-sensor fusion considering attention weights and fault pattern-based ranking and addressed both long-term and local dependencies with general and recurrent transformer models.

Chapter 7 concludes the research work and provides future directives to extend the work to more challenging dimensions in the future.