mathematics

Review

Prognostics and Health Management of Rotating Machinery of Industrial Robot with Deep Learning Applications—A Review

Prashant Kumar , **Salman Khalid** and **Heung Soo Kim** *

Department of Mechanical, Robotics and Energy Engineering, Dongguk University-Seoul, Seoul 04620, Republic of Korea; prashantkumar@dgu.ac.kr (P.K.); salmankhalid@dgu.ac.kr (S.K.)
* Correspondence: heungsoo@dgu.edu; Tel.: +82-2-2260-8577; Fax: +82-2-2263-9379

Abstract: The availability of computational power in the domain of Prognostics and Health Management (PHM) with deep learning (DL) applications has attracted researchers worldwide. Industrial robots are the prime mover of modern industry. Industrial robots comprise multiple forms of rotating machinery, like servo motors and numerous gears. Thus, the PHM of the rotating components of industrial robots is crucial to minimize the downtime in the industries. In recent times, deep learning has proved its mettle in different areas, like bio-medical, image recognition, speech recognition, and many more. PHM with DL applications is a rapidly growing field. It has helped achieve a better understanding of the different condition monitoring signals, like vibration, current, temperature, acoustic emission, partial discharge, and pressure. Most current review articles are component- (or system-) specific and have not been updated to reflect the new deep learning approaches. Also, a unified review paper for PHM strategies for industrial robots and their rotating machinery with DL applications has not previously been presented. This paper presents a review of the PHM strategies with various DL algorithms for industrial robots and rotating machinery, along with brief theoretical aspects of the algorithms. This paper presents a trend of the up-to-date advancements in PHM approaches using DL algorithms. Also, the restrictions and challenges associated with the available PHM approaches are discussed, paving the way for future studies.

Keywords: prognostics and health management (PHM); deep learning (DL); industrial robots; rotating machinery

MSC: 68T01

Citation: Kumar, P.; Khalid, S.; Kim, H.S. Prognostics and Health Management of Rotating Machinery of Industrial Robot with Deep Learning Applications—A Review. *Mathematics* **2023**, *11*, 3008. https://doi.org/10.3390/math11133008

Academic Editor: Faheim Sufi

Received: 23 June 2023
Revised: 3 July 2023
Accepted: 4 July 2023
Published: 6 July 2023

1. Introduction

Industrial robots (IRs) have drawn a lot of attention over the past two decades due to the availability of cutting-edge technologies and the need for high production. Industrial robots have found applications in almost every sector, like manufacturing, underwater exploration, hazardous material disposal, steel exploration, and entertainment [1–5]. IRs include multiple forms of rotating machinery, like servo motors and gears, which are also prone to failures [6]. In general, IRs are robust machines; however, faults are inevitable. The rotating machines are vital components of the IRs and act as a driving force [7]. IRs have seen a surge in their applications in recent years [1,8–14]. This has aided human efforts to reduce the operational costs. The World Robotics 2021 industrial reports show that more than three million IRs are operating globally, almost 10% more than the previous year. In 2020, despite the global epidemic, new robot sales climbed by 0.5%, with 384,000 units shipped worldwide [15]. With a compound annual growth rate (CAGR) of 11.7% between 2021 and 2030, the size of the worldwide IRs market is predicted to increase from USD 37,876.0 million in 2020 to USD 116,848.7 million by 2030 [15]. In 2021, the size of the global IRs market was estimated at USD 15.60 billion [16]. The market for IRs was estimated to be worth USD 26.52 billion in 2022 and is anticipated to grow at a CAGR of 10.5% between

2023 and 2030 [17]. The global trend of publications containing the keywords "industrial robot fault" in the title that were published per year, as determined by the Web of Science and PubMed, is given in Figure 1.

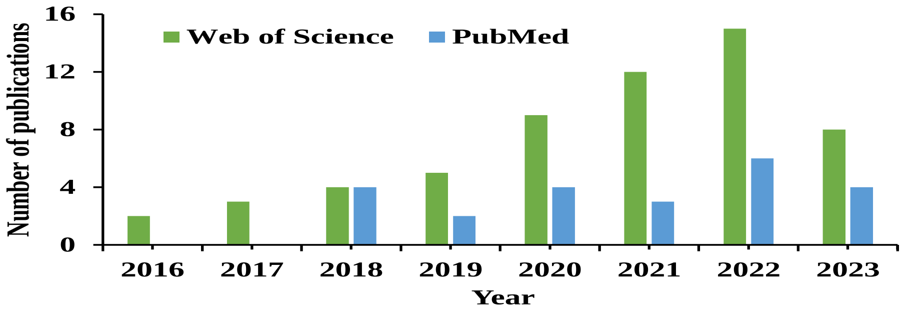

Figure 1. Global trend of the publications containing the Keyword "industrial robot fault" in the title that were published per year, as determined by the Web of Science and PubMed.

The market trend shows that IRs will be a driving force in industry. The robot system applications in those industries help to improve productivity, efficiency, and quality. The technological improvement in sensors, motors, and drives has improved the performance and efficiency of IRs. Due to advancements in robot system position and trajectory precision, arc welding has gained popularity over spot welding in various applications.

The technologies involved in robots have become complex and require a lot of feedback from the environment for efficient operation and precise control. The reliability of robots and their associated components is critical for minimum downtime and maximum production [18]. Robotic system health monitoring, diagnostics, prognostics, and maintenance have received a lot of attention as a result of the high-reliability requirements. The PHM of advanced robotics setup in industries are crucial for the smooth functioning of production and serviceable units. An efficient PHM approach for industrial robots and their rotating machines is the need of the hour [19]. A holistic framework for PHM [20–24] is shown in Figure 2. It comprises the data collection, data conditioning, fault detection (FDT), fault diagnosis (FDG), fault prognosis (FP), and decision support in a chronological manner [25–28]. The domains of fault detection, diagnosis, and prognosis have been extensively studied. The focus of this paper will be restricted to these topics.

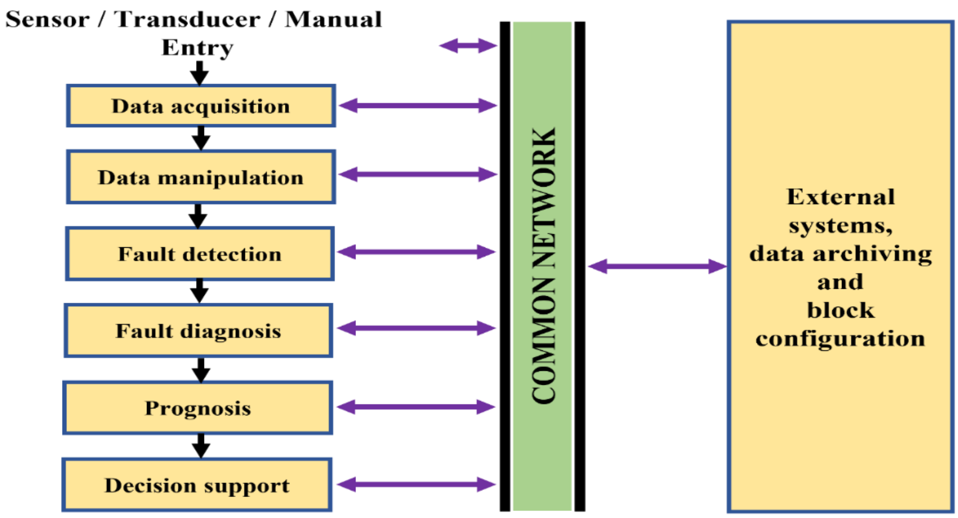

Figure 2. Holistic framework of PHM.

PHM is a modern engineering strategy that amalgamates advanced sensing methodologies, failure physics, statistical analysis, artificial intelligence (AI), and reliability analysis to enable the real-time health monitoring and prognostics of a system's future state based on the existing data [27]. PHM refers to a collection of strategies and processes for monitoring, diagnosing, prognosticating, and maintaining a machine or process [29]. Manufacturing systems use PHM technologies to reduce unplanned downtime and expenses. PHM aids engineers in transforming data into interpretable information and assessing health, improving the system's understanding. It makes it possible to create strategies for the system's efficient and planned operation. Initially, aerospace industries used PHM strategies, but they have also found applications in several fields, like manufacturing, automobile, and railways [9–13]. The ability to estimate a system's remaining useful life (RUL) through PHM while it is in use makes condition-based maintenance (CBM) possible. It helps develop a maintenance strategy in which only damaged parts are repaired or replaced [30–34]. CBM is a systematic approach that combines hardware and software to continuously assess the equipment performance and deterioration without interfering with the system's normal operations [35,36]. CBM uses the actual condition of the equipment as opposed to system/component breakdowns or planned maintenance. Prognostics is a vital element for CBM as it enables timely maintenance decisions [37–39]. The concept of preventive maintenance leads to a rise in expenses for many industrial companies, as many components are replaced before the end of their lifecycle. Therefore, maintenance should be carried out as needed to ensure a high level of safety and dependability. This is the core concept of CBM, and PHM is the key technology for realizing it.

At present, diagnostics is conducted with the help of instruments, like sensors, meters, controllers, and computational devices [40]. These devices are used to obtain signals from the machine or process for diagnostic purposes. The root causes of failure can be identified using sophisticated diagnostic approaches [41,42]. The diagnostic task is a reactive maintenance process that is performed when a fault actually occurs. The standalone application of diagnostic approaches does not have a significant effect on reducing the occurrence of downtime and the related expenses. To improve the management of the maintenance scheduling and production optimization, maintenance should be conducted in a proactive manner [43]. This can be accomplished by switching maintenance strategies from typical break-and-repair (diagnostics) to predict-and-avoid (prognostics). The aim of PHM is to establish and deliver an integrated strategy for viewing the machine's health to users. PHM involves both prognostics and diagnostics [44]. By identifying and establishing the causal connection between cause and effect, diagnostics is the method of discovering defects and identifying the primary causes of failure. The practice of evaluating and predicting health, which includes anticipating an impending failure and the remaining usable life, is known as prognostics [17–19]. Implementing timely and suitable maintenance actions and making precise logistics decisions based on the diagnostic and prognostic outputs, available resources, and operational demand are all parts of health management.

With the advancement in sensor technology and computational power, artificial intelligence (AI) has attracted researchers to improving the existing PHM approaches, as well as developing new methodologies. Different AI algorithms, like support vector machine (SVM), random forest (RF), k-nearest neighbor (kNN), decision trees (DT), artificial neural network (ANN), and many more, have been used for fault diagnosis (FD) and fault prognosis (FP) [44]. FD and FP methodologies using these AI algorithms require suitable features as input, which requires prior knowledge and expertise of the fault. This creates a hindrance in developing PHM solutions with generalization capabilities. The availability of cloud computing, huge data storage capabilities, sensors, communication technologies, and a complex engineered setup have led to huge data generation and collection [20–22,45–50]. Important details regarding the condition of the system are provided by this data. The development of multidimensional and heterogeneous data streams has a tremendous impact on the operation of traditional AI methods, like SVM, kNN, RF, and DT [51–53]. More refined analytical tools and improved approaches are required to efficiently and inherently harvest the features concealed in actual-time measured systems.

Over the last decade, deep learning (DL) has attracted researchers from different domains like biomedical, image recognition, natural language processing, and voice recognition systems worldwide, owing to its excellent properties [23–28]. A deep learning algorithm has immense potential. Deep networks are used to spontaneously manage extremely non-linear and sophisticated feature extraction from unprocessed information, eliminating the necessity for manual feature development [54–60]. DL can spontaneously discover hierarchical features from enormous and multidimensional industrial data, making it a viable tool for the PHM solution [61–64]. Lee et al. [65] have proposed a fault detection approach for the robotic servo-motor under varying working conditions. Rauf et al. [66] have proposed a transfer learning-based DL approach for fault detection in the industrial robotic system. Zhou et al. [67] have proposed a harmonic reducer fault diagnosis using the deep learning-based model. Adam et al. [68] developed a multiple fault diagnosis approach with the help of a convolutional neural network-based algorithm. Yin et al. [69] have developed a dual-driven transfer network for fault diagnosis in industrial robots. Figure 3 shows the various ways in which PHM strategies can be developed. The task of feature extraction and selection is primarily emphasized in traditional data-driven techniques. It is heavily reliant on signal processing techniques and human knowledge. These approaches require numerous adjustments when working with big volumes of data and do not operate in real time. DL models are capable of automatically discovering and removing pertinent features from unprocessed data. By doing so, manual feature engineering—which can be time-consuming and prone to mistakes in fault detection tasks—is no longer necessary. DL models have the capacity to immediately learn intricate patterns and representations from the data, improving defect detection. DL models can effectively handle complicated datasets. They have the capacity to learn from a variety of data sources and identify subtle patterns that conventional approaches can find challenging. DL models can handle complex fault patterns and nonlinear interactions. The nonlinear behaviors of many industrial systems can be difficult to model using conventional techniques. End-to-end learning, where the model learns directly from the input data to the output predictions, is made possible by DL models. As a result, manual intervention at crucial points in the pipeline for fault detection is no longer required. In a nutshell, DL provides an end-to-end framework and facilitates a unified PHM system.

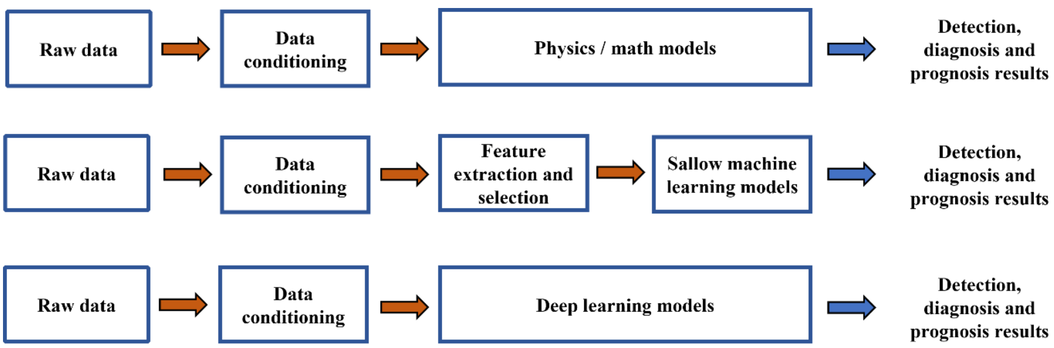

Figure 3. Various PHM methods with Physics/math models, ML-based model, and DL-based models.

Most current review articles are component- (or system-)specific, and do not reflect the advancement in DL-based strategies. This is a burgeoning field, and several studies are being conducted to develop more refined approaches and improved strategies. Many advanced methods are emerging each month, and there is a need to review the latest trends and PHM strategies. Few papers have reviewed the DL-based PHM strategies for rotating components of the IRs. This paper examines the PHM strategies based on the DL-based approaches for industrial robots and their rotating machinery. Section 2 discusses faults and failures in industrial robots and their rotating components. Section 3 illustrates the PHM methodologies, together with the conventional PHM cycle, and details the PHM performance metrics and DL-based PHM. Section 4 presents a brief description of the various DL algorithms. Section 5 presents a detailed study of the existing DL framework-based PHM for the rotating machinery of the IRs. Section 6 briefly discusses the PHM strategies with DL applications and future possibilities. Finally, Section 7 concludes the work.

2. Industrial Robot Configuration and Faults

Industrial robots are complex and require continuous monitoring for optimum performance and minimum downtime. A robot is described by the International Standards Organisation (ISO 8373) as a machine that has automatic control, reprogramming, and manipulation capabilities [70]. This machine is flexible and has numerous configurable axes. It is made to be used in a variety of industrial applications and can be either stationary or mobile. An IR is a general-purpose programmable machine that possesses the characteristics of the human arm. It can be programmed by its computer to move its arm through sequences of motion to perform some useful tasks. It can perform a similar motion over and over until it is reprogrammed to perform other functions. Many industrial operations involve robots working together with other equipment. The main features of the robot include:

- A robot can produce a job with consistent quality at a steady state with practically zero rework and wastage.
- Robots can work continuously throughout the work cycle with proper maintenance solutions.
- Robots' upkeeping cost is increasing at a lower price in comparison to the labor maintenance cost every year.
- The capital cost for the robot is paid once only.
- Robots can take up repeated tasks and challenging jobs even in an unsafe and unhealthy environment.
- Robots can work precisely at higher speeds and can exert larger force than in humanly possible.

IR comprises different components, like robotic arms, body, arm, actuators, rotate vector reducers, sensors, end-effectors, switches, gears, and linkages. There can be numerous faults in a robotics system due to its complex nature. Figure 4 shows a block diagram of the robot system. Three essential parts make up a robotic system: the power sources, the computer used to manage the robot, and the robot's mechanical framework. All necessary pneumatic, hydraulic, and electro-mechanical components are included in the robot's mechanical design. This includes electrical actuators, which are motors used for rotational operations, as well as non-electric actuators that utilize hydraulic or pneumatic systems, or both. These components collectively enable the robot to carry out its intended functions. The robot has many internal sensors, which are mainly used for measuring the rotary positions of the motor shafts, gears to reduce the speed between the motors and the joints, switches, and relays for creating selected operations. These motions will have to be performed when certain conditions are met. The robot has an end effector, such as a tool or a gripper. The entire mechanical structure is interfaced with the robot control computer. The robotic system's computer comes equipped with a variety of software applications required for the structure of the robot to function. These software packages incorporate coordinated transformation software, which makes it easier for the robot's movements to be seamlessly coordinated. To control the actuators' speed and location, control software is also present. Additionally, the computer has interfaces for teaching and learning particular tasks, enabling users to guide the robot successfully. Additionally, it includes safety precautions to guard against any potential harm to the robot structure, thus creating a safe working environment. Figure 5 shows a pictorial view of the IR, while Figure 6 depicts the faults in the reducers of the IR. Figure 5 demonstrates the six degrees of freedom (DOF) IR. The six DOF means that the 6-axis of the IR can move independently. There are several types of IRs, which include the non-servo robot, servo robot, programable robot, and computer programmable robot. A non-servo robot is typically used for moving objects, like picking up an object and transporting it to another place. The manipulators and effectors—robotic appendages that serve as the robot's arms and hands and provide it enhanced flexibility and greater movement—enable the servo robots to perform a range of tasks. A programmable IR can execute the repeating task a certain number of times based on fed programs. A servo robot that can be programmed by a computer and controlled remotely is known as a computer-programmable robot. Also, IRs have many structural configurations that suit several applications in the industry.

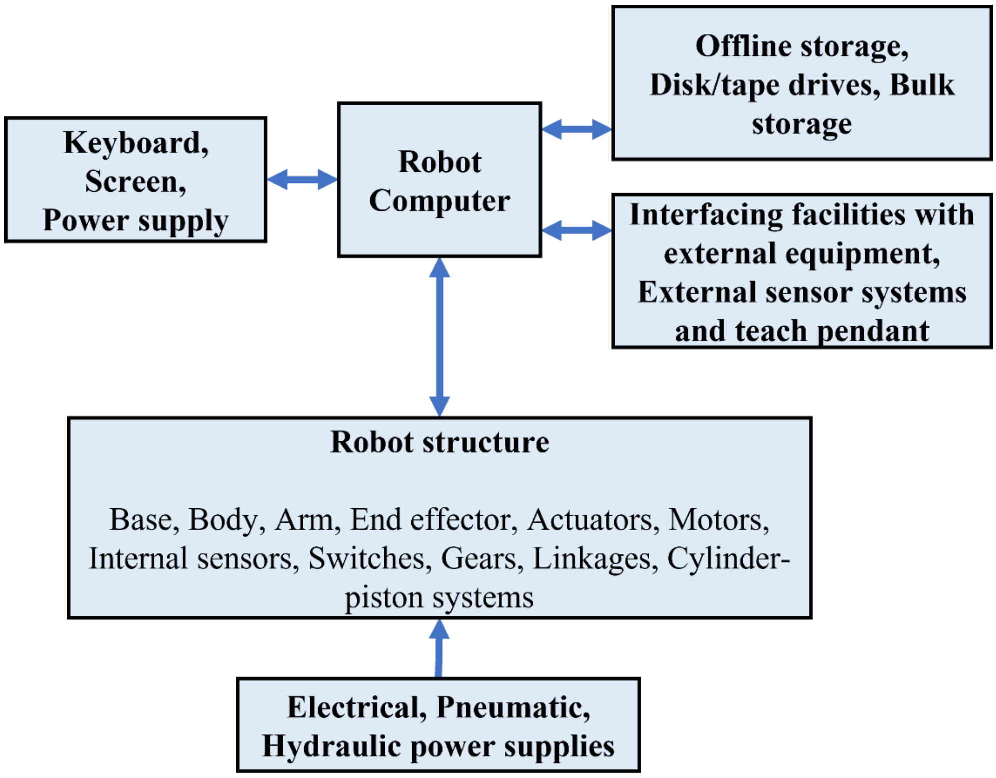

Figure 4. Block diagram of IR system.

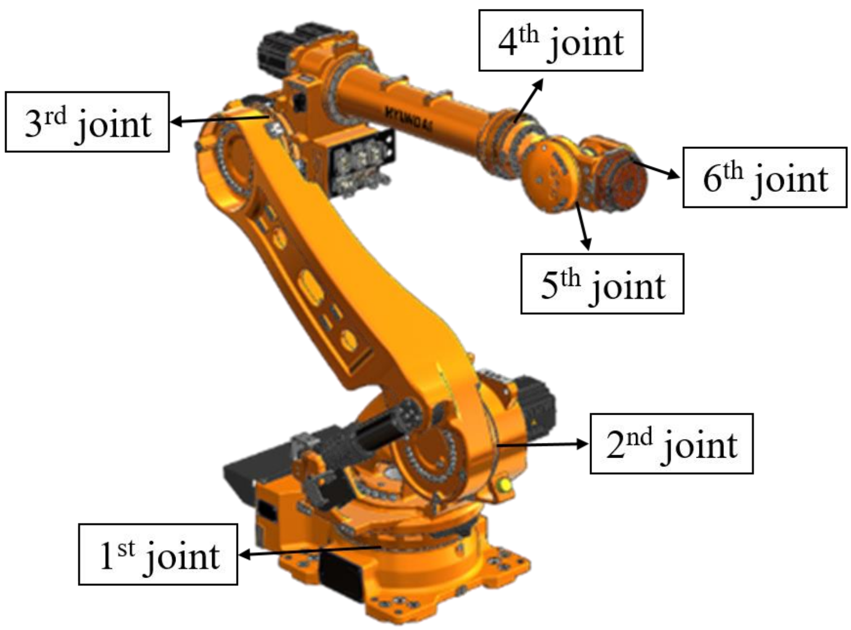

Figure 5. Industrial robot with 6 degrees-of-freedom [48].

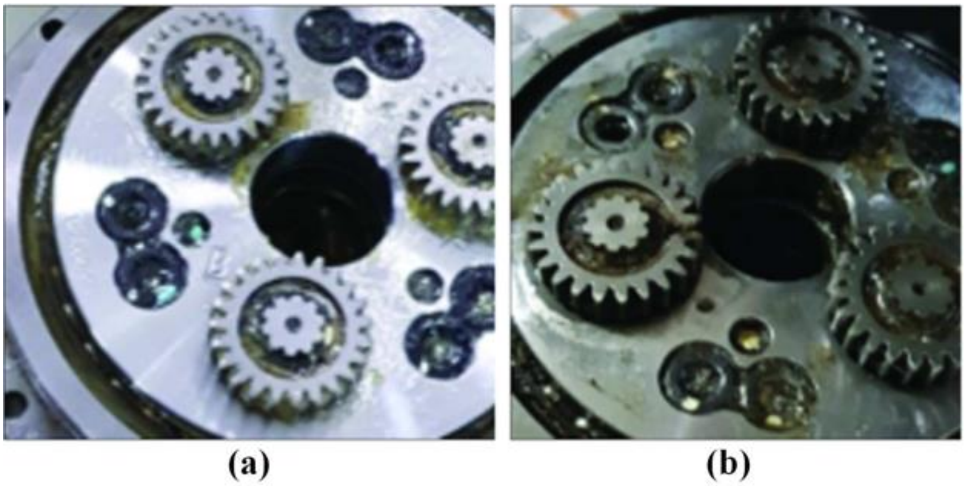

Figure 6. Faults in the industrial robots (**a**) Faulty reducer, and (**b**) Faulty Aged reducer [48].

It is vital to detect faults in a timely manner, and research on the PHM of the robotic system could be vital for efficient FDT. Faults, soft failures, and hard failures are the three types of faults that can be found. A fault is a systemic issue, such as a defect, an inaccurate signal value, or a poor decision. A fault may cause the degradation of the system's performance and can lead to failure [71]. Faults in the system can cause the system to wear out faster. The system may nevertheless be able to achieve the desired productivity and output product quality. A soft failure is characterized by deterioration, 'wear and tear', and exterior variations, which result in system damage. A soft failure occurs when a system continues to function, but its performance begins to deteriorate. In the case of a soft failure, the system's productivity decreases, and it is unable to meet the required objective. With time, its productivity decreases and its performance worsens, and ultimately decreases to a point below the required standards. A hard failure can be referred to as a breakdown of the component/equipment of the system, which leads to a pause in the functioning of the system [71]. Under such conditions, the manufacturing process is compromised, and cannot meet the required demand. The whole process completely shuts down. The system's PHM should be capable of detecting both soft and hard failures. PHM should be designed in such a way that it can detect failures early enough to prevent significant breakdowns. The majority of PHM solutions are centered on component monitoring and failure. There is a necessity for system-level monitoring, which will be applied to track the source of the failure to the point of genesis, or wherever it may have originated, by monitoring the health of the components, as well as the system.

The rotating components of the IR are crucial as well as being vulnerable to the faults. The rotating components, such as the motors, gears, and reducers, are also vulnerable to faults. The PHM of these rotating components could significantly reduce downtime and help in scheduling maintenance. The faults in the rotating components include faults in the gearboxes, centrifugal pumps, motors, alternators, and many more. The faults in the rotating machinery can include bearing defects, rotor defects, eccentricity, gear wear, cavitation, and misalignment [72–74]. Bearing defects can be segregated into outer race defects, inner race defects, ball defects, and train defects. Eccentricity can be divided into static, dynamic, and mixed eccentricity. Rotor defects are the major issue in the induction motors, alternators, and generators. High starting torque and frequency switching are the major causes of rotor damage in the motors [75]. The symptoms of the faults can include high vibration, increased current demand, torque pulsation, and excessive heating [75]. Timely maintenance of the rotating machinery is essential to avoid a complete shutdown. The PHM of rotating machines is critical in achieving the aim of uninterrupted operations in industries.

3. PHM Methodologies

PHM has previously existed in the medical and aerospace fields. PHM involves three subfields; namely, fault detection, fault diagnosis, and fault prognosis [76]. FDT aims to detect instances when the machine starts to behave differently from its normal behavior. It can be treated as a dual categorization job, i.e., to categorize whether the machine is running well or unwell [77]. Fault diagnosis involves fault identification, fault localization, and recognition of the severity. The next phase, diagnostics, should be able to pinpoint what went wrong and should build on the understanding that something went wrong. The analysis and prediction of the fault diagnosis should be more thorough than that of FDT. In the fault prognosis, the RUL is estimated. Prognostic models and physics-based models are typically utilized. The RUL is estimated with the help of the degradation trajectory [78]. From a practical perspective, the correct RUL estimation is crucial, as an incorrect estimation will lead to over-maintenance or complete shutdown. An accurate RUL assessment will help in adequate maintenance scheduling. The PHM solution combines these three groups to provide an optimized solution to the machine's maintenance. FDT application through applying DL can be categorized into two groups: supervised and unsupervised [79,80]. Supervised learning involves the availability of labelled data in the training and test dataset. Any DL approach can be chosen depending on the nature and availability of the data. Fault diagnosis can be viewed from the perspective of AI as a multi-class classification problem. It entails categorizing the detected fault according to a particular set of fault type, location, and severity. As the target value is in the actual world, FP in AI applications may be viewed as a regression problem [81]. The prediction seeks to create a learning function that could translate the state of the machine to its RUL [82–85].

The variety of the sensors data are used for the PHM of IRs. These sensors data help in assessing the health of the IRs. Accelerometers, encoders, temperature sensors, current and voltage sensors, vision systems, and force/torque sensors are a few of the frequently utilized sensors. The vibration sensor data have been significantly used for the PHM of IRs. The joints and actuators of the robot are measured in terms of their position, velocity, and direction using encoders. The system can find differences between the actual position and the expected position, which may point to faults with the robot's mobility or control system. Robotic joints, motors, electronics, and other parts are all monitored for their temperature using temperature sensors. Overheating or other potential defects can be indicated by sudden temperature fluctuations or by exceeding the predetermined thresholds. The electrical parameters of the robot's motors and actuators are monitored using current and voltage sensors. Variations in the current or voltage levels may be a sign of electrical problems, such as overloading, short circuits, or other issues. The forces and torques applied by the robot during its interactions with the environment are measured by force and torque sensors. These sensors can identify anomalies, such as sudden contact forces or high torques, which could be signs of an impending collision or malfunction. Robotic vision systems are used to keep an eye on its surroundings. These systems include cameras and image processing software. The system can discover possible issues through their ability to spot visual irregularities such as erroneous part arrangement, missing objects, or variations from the typical visual patterns.

For a long time, vibration-based analysis has been widely utilized for prognostics due to its superior capabilities, and many applications still employ this traditional method [37–40]. The vibration data have been widely used for developing a PHM strategy in rotating machines. It is one of the PHM topics that has been the subject of the most research. Other techniques, like acoustic emission, temperature analysis, and ultrasonic, are also widely used. Processing is conducted on the sensor data and is amalgamated using the sensor fusion techniques, owing to their innate advantages [86–90]. Model-based, data-driven, and hybrid prognostic techniques are the three types of prognostic technologies currently available [91].

A model is created and simulated for the healthy and fault states in a typical model-based process. The assessment of the developed model under the system's many functional modes is used to approximate the system's remaining useful life (RUL). This is created by combining the time-averaged model probability and the weighted predictions from each mode. The reliability of many models is under scrutiny, and if it is unavailable, data-driven techniques are utilized to estimate the RUL. This is generally accomplished by visualizing the developing fault's trajectory and the time it takes to obtain to the preset threshold value. The two famous fault prediction tools are the Kalman filter and the Alpha–Beta–Gamma tracking filter, which are often used in aerospace PHM and many other fields [43–47,92]. In hybrid approaches, both model-based techniques and data-driven methods are amalgamated and used for fault prognosis and diagnosis [93–97]. The traditional method of PHM is based on analytical models that draw on physical principles and subject-matter expertise. Systems with well-understood physics and well-defined failure mechanisms may respond well to this strategy. Analytical models for complex systems can be difficult and time-consuming to create, and they may not be correct in the presence of unexpected or unpredictable behavior. Machine learning algorithms that learn to recognize patterns in the data provide the foundation of the DL approach to PHM. This method can be used to analyze complicated systems with a wide range of failure modes as it is not constrained by the requirement to comprehend the underlying physics of the system. Deep learning algorithms are more resistant to unforeseen or unpredictable behavior because they may learn to adapt to changes in the system.

An efficient PHM strategy should detect incipient faults, diagnose faults, and estimate the RUL of the component or sub-elements. Both products and processes can benefit from PHM. The focus of product PHM is on a physical object. Monitoring a robot arm is an example of product PHM. In comparison to process PHM, product PHM is more readily available in the automobile, aerospace, and power generation industries [49–56].

3.1. Conventional PHM Cycle

Prognostics and Health Management comprise multiple tasks to lower the overall lifecycle cost of the component/system. The PHM strategies involve multiple steps (as shown in Figure 7), like data collection, feature development, dimensionality reduction, model development, decision making, and remaining useful life calculation. The data acquisition step involves collecting the data, like the vibration and current temperature, from multiple sensors, including the accelerometers, acoustic emission sensors, thermometers, and hall sensors. These data contain information regarding the health of the machine. The feature extraction step involves the application of signal processing tools, like fast Fourier transform (FFT), Short-time Fourier transform (STFT), Wavelet packet transform (WPT), and Hilbert Huang transformation (HHT). The statistical features based on time-domain signals are kurtosis, root mean square (RMS), skewness, etc., and are used to develop PHM strategies [98–108]. Also, some of the frequency-domain signatures, like spectral, envelope, and wavelet packets, are widely employed for PHM strategies [20,57–60]. The feature selection step involves removing redundant and irrelevant features and is accomplished by selecting the essential features using filters, wrappers, or embedded methods. Also, dimensionality-reduction tools, like principal component analysis (PCA), linear discriminant analysis (LDA), and kernel PCA, are used for feature dimensionality reduction, as well as for retaining rich information about the health of the intact machines [109,110].

The traditional method for anomaly detection includes SVM, Hidden Markov models, the Bayesian network, and ensemble methods. These methods have been applied efficiently for the health assessment of different machines [111,112]. PHM involves three tasks: diagnostics, prognostics, and decision support. Diagnostics is the crucial task following FDT to understand the system's health by analyzing the severity level of faults. The conventional machine learning approaches involve the SVM, kNN, DT, and RF trained on a labeled dataset for fault identification and classification [111,113,114]. Prognostics refer to the detection of incipient faults and related RUL for the predictive maintenance of the system. The data-driven approaches, like ANN, HMM, the Kalman filter, and the extended Kalman filter, have been utilized for the prognosis [43,67–69]. The PHM strategy's health management system is referred to as decision support, which utilizes the outcome of the diagnostics and prognostics for making timely, suitable, and logical judgments to schedule the maintenance or replacement of components [115–121]. To determine the best maintenance task and time to apply it, mathematical programming, Markov decision processes, and Reinforcement learning (RL) methods are prominently used [71–73].

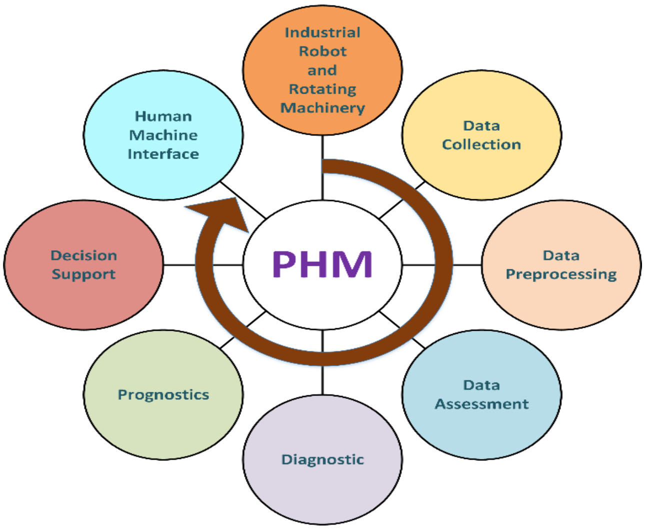

Figure 7. Framework for the conventional PHM.

3.2. PHM Performance Metrics

The performance analysis of the PHM approach determines the system's reliability. The complexity of the PHM system requires the appropriate performance metrics, which are listed in Table 1. These metrics are employed to assess the RUL prediction for prognostics. The performance metrics offer a thorough examination of how well the PHM techniques function with DL applications.

Table 1. Performance metrics for PHM evaluation.

Diagnostics	Prognostics
Accuracy [122]	Mean absolute error [123]
Error rate [122]	Root mean square error [124]
Precision [125]	Mean absolute percentage error [124]
Sensitivity [126] (p. 202)	Prediction horizon [127]
F1-score [126]	Convergence [127]
Correlation coefficient [128]	Relative accuracy [127]
Area under curve [129]	Confidence interval [130]
Detection error trade off [129]	Exponential transformed accuracy [131]

3.3. DL-Based PHM

The application of DL in PHM has gained momentum in the last few years. With the inherent capabilities of DL-based models, the disadvantages of conventional ML-based models have been seized. The DL algorithms like CNN, RNN, AE, etc., offer automatic feature development, which significantly improves the model's performance. The DL-based model has been efficiently used for FDT, FDG, and FP. All of the major PHM fields have a universal framework as a result of DL. This can be illustrated by the simple diagram shown in Figure 8. The application of DL provides an end-to-end learning framework for PHM. The type of data available and the application domain influence the choice of the DL model. When there is a scarcity of labeled data, which often happens in practical problems, FDT requires unsupervised learning. A multi-class classification challenge could be said to exist in the fault diagnosis scenario. The objective of the DL created for fault diagnosis is to map the detected fault to a certain combination of fault type, location, and severity. A typical DL model will involve the SoftMax layer to the final layer for achieving the fault diagnosis task. A common choice for the loss function is categorical cross-entropy. The model is trained based on this loss function. Also, after training a DL model, the t-SNE method can be used for feature visualization. The prognosis task can be considered as the regression task. The RUL prediction might be reduced to a normalized range by the final layer of the DL model, which could be a single neuron with a linear activation function or sigmoid function. The accurate estimation of the RUL is a grueling task as an overestimation could lead to unnecessary maintenance and an underestimation could lead to a complete shutdown of the machine. One of the most important tasks in the prognosis is penalizing the delayed RUL estimations (i.e., the predicted RUL is higher than the actual RUL). The input data for the PHM solution for robots and their rotating machinery can include vibration data, current data, imagery data, temperature data, etc.

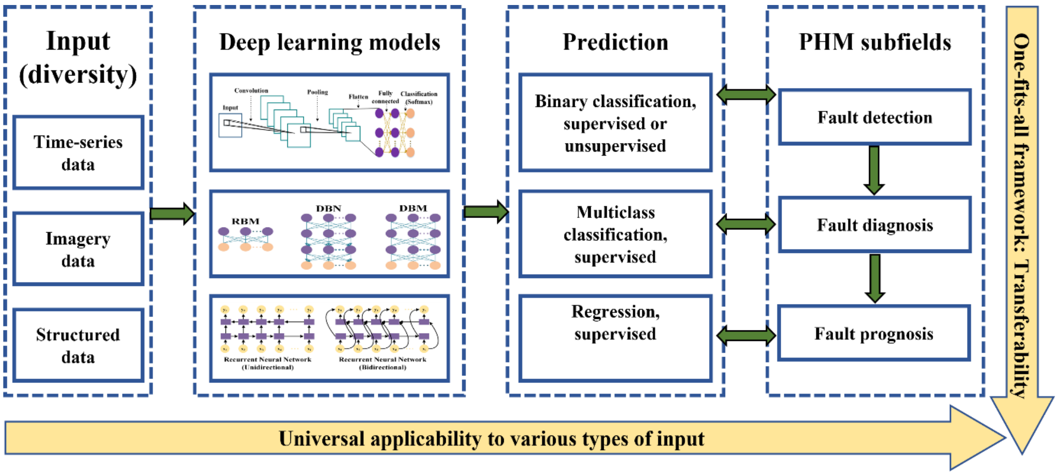

Figure 8. DL-based framework for the PHM including its subfield like fault detection, fault diagnosis, and fault prognosis for various types of input.

Deep learning framework-based PHM systems often incorporate different deep learning architectures, including deep Boltzmann machines, autoencoders (AEs), convolutional neural networks (CNNs), and recurrent neural networks (RNNs). An in-depth explanation of how these designs are applied in PHM systems is provided below:

- Restricted Boltzmann Machines (RBMs): RBMs have a wide range of applications; their direct use in PHM systems built on deep learning frameworks has been relatively less common. RBMs, however, can contribute to PHM in many ways. In PHM, RBMs can be utilized to find anomalies. An RBM can learn the underlying distribution of the normal behavior by being trained on data from typical operational scenarios [132]. The RBM may assess the reconstruction error or energy of fresh data instances during the inference stage. Higher reconstruction errors or energies signify abnormalities or flaws because they deviate from the expected behavior. In PHM systems, RBMs can be used as a step in the pre-processing pipeline. RBMs are capable of extracting features from high-dimensional sensor input or learning a compressed representation. The hidden units of an RBM can be trained on the input data to identify significant latent characteristics or patterns that can be used as inputs to later models, such as fault classifiers or prognostics models [133].
- Autoencoders (AEs): Autoencoders are able to pick up on a system or component equipment's typical working behavior and recognize abnormalities or departures from it. Autoencoders identify probable errors or anomalies by highlighting variations between the original and reconstructed input data [134]. In PHM, AEs can serve as feature extractors. The encoder portion of an autoencoder can capture meaningful representations of the sensor data by being trained on a sizable dataset. These rep-

resentations can then be utilized as inputs for later supervised models, such as fault classifiers or prognostics models [135]. High-dimensional sensor data can have its dimensions reduced by AEs, allowing for more effective processing and storage. AEs preserve important information while simplifying later modelling efforts by lowering the dimensions of the input they compress into a lower-dimensional latent space [136].

- Convolutional Neural Networks (CNNs): The capacity of CNNs to efficiently extract spatial patterns and features from the sensor data, such as images or time-series data, makes them a common tool in PHM. CNNs can be used for fault detection or classification tasks in applications where images or visual data are accessible (such as thermal imaging or photos from visual inspection) [137]. CNNs develop hierarchical representations of the images they process, identifying pertinent details and patterns linked to errors or anomalies [138]. It can be used to analyze spectrogram data, which displays the frequency content of time-series sensor measurements in signal-based PHM. In order to perform tasks like fault detection, classification, or regression, CNNs may extract spatial patterns from spectrograms [139].
- Recurrent Neural Networks (RNNs): RNNs are made to identify sequential patterns and temporal dependencies in time-series data. Sequential sensor measurements are frequently used in PHM applications, making RNNs an excellent choice for this type of data analysis [140]. RNNs can simulate temporal dependencies in time-series sensor data, especially those with Long Short-Term Memory (LSTM) or Gated Recurrent Unit (GRU) variations. Using RNNs, it is possible to detect faults and perform diagnostics and prognostics on sequential data by capturing patterns, long-term dependencies, and dynamics [141]. In PHM, time-series forecasting tasks can be performed using RNNs. RNNs can forecast future sensor readings, remaining usable life (RUL), or failure probabilities by learning from the existing sensor data, which enables proactive maintenance planning [142].

4. Overview of Deep Learning Models

Deep learning has gained popularity owing to the availability of high computational resources. The idea of DL dates back to the 1940s [143], but appears to be a new concept, as it was relatively unknown for several years before gaining traction, and it was known by a variety of names prior to being termed “deep learning”. DL developed in three stages: cybernetics in the 1940s–1960s, connectionism in the 1980s–1990s, and the present revival under the name DL, which began in 2006 [143]. Deep networks are based on the human brain’s hierarchical architecture and attempt to learn simple patterns; they transform them into more abstract representations [144–146]. The generic structure of a feed-forward deep neural network (DNN) contains an input layer, numerous hidden layers, and an output layer. When multi-layer perceptron (MLP) obtains the input data, the output is generated along with the successive layers of the model in a straightforward manner. The non-linear activation functions of each hidden/middle layer neuron are given the biased weight sum of the preceding layer outputs to generate the neuron’s output. The DL model’s hierarchical design allows for efficient feature learning, which aids in comprehending the underlying correlations and patterns in enormous amounts of data [147,148]. The following section briefly describes the available DL algorithms that are applied in the PHM strategies.

4.1. Restricted Boltzmann Machine

The Restricted Boltzmann machine (RBM) is a generative stochastic neural network framework. It can discover a probability distribution within a collection of inputs. The RBMs are undirected bipartite graphical models with n_x visible and n_h hidden units that allow no intralayer connections and are widely employed as generative models (GMs). Due to their stochastic processing units, RBMs can learn the original data’s recreated form. They are generally used as a pre-processor for various frameworks to complete the job in supervised classification, but can also be utilized as a standalone classifier. Goodfellow et al. [143] provides a step-by-step strategy for training RBMs. In the coming sections, we briefly discuss two generative deep neural network (DNN) models based on the RBM; namely, the deep belief network (DBN), and deep Boltzmann machines (DBMs). Figure 9 shows the structures of the RBM, DBN, and DBM.

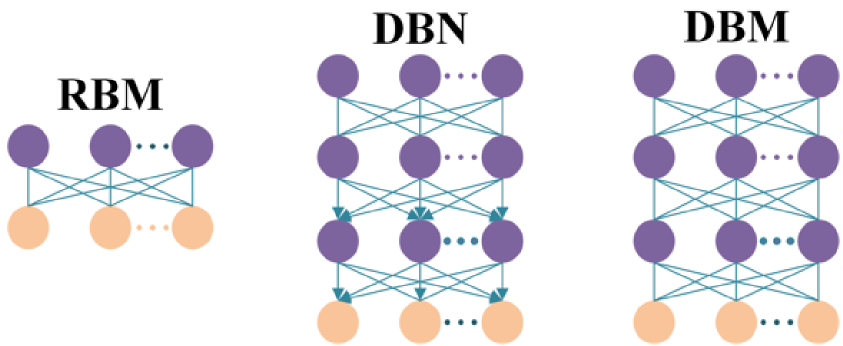

Figure 9. Structure of the RBM, DBN, and DBM (shaded boxes in lavender color represent hidden units).

4.1.1. Deep Belief Network

The DBN is a deep probabilistic GM constructed by stacking several RBMs, and consists of several layers of stochastic, latent variables [149]. Hidden units or feature detectors are terms used to describe latent variables that have binary values. The two layers at the top are linked by undirected and symmetric connections, which form an associative memory. The upper layer sends directed links down to the lower layers. The states of the units produce the data vector in the lowest tier. Links in the lower layers are top-down directed, whereas the top layers are undirected. An effective layer-by-layer process is utilized to learn the generative weights. These weights specify the relationship between the variables in different layers. After learning, a single bottom–up run is performed. It begins with the help of a detected data vector in the end layer. To determine the values of the latent variables in each layer, the weights generation process is reversed. The nets of DBN perform one layer of learning at a time when inferring the data. This is conducted by employing the latent variable values of one layer as the training data for the coming layer. It can also be amalgamated with other learning methods. It can help in fine-tuning all the weights that will boost the multiplicative or discriminatory operation of the complete framework.

4.1.2. Deep Boltzmann Machine

The Deep Boltzmann machine is a deep GM with layers that are organized hierarchically. The DBM is formed by stacking layers of the RBM, such that odd-numbered layer units and even-numbered layer units are independent. DBM is a totally undirected model, unlike the deep belief network. The latent variable in a DBM contains numerous layers, whereas RBMs only have one. Within each layer of a DBM, the variables are mutually independent and conditioned on the variables of the neighboring levels. The DBM's energy function incorporates a weight matrices-based connection between hidden units (latent variables). A DBM can also be arranged in a bipartite graph. In contrast to a DBN, which may be taught layer-by-layer, a DBM is trained as a joint model. As a result, DBM training is more computationally expensive than DBN training.

4.2. Auto-Encoder

An auto-encoder (AE) is an efficient neural network model that employs an unsupervised learning method for learning efficient data coding in an unsupervised way. It is made up of two parts: an encoder and a decoder. The encoder is utilized to encode the input, and in some cases, to compress the data. Each layer of the encoder has a decreasing number of hidden units. It allows only the most significant and representative attributes to be mined from the data. The decoding part of the framework is the second component. Each layer of the decoder has an increasing number of hidden units, and the decoder tries reconstructing the original input using the encoded data. Consider the encoder, which uses a non-linear mapping to turn the input i into a hidden representation e, and given by [143,150]:

$$e = \phi(W * i + b) \tag{1}$$

where, ϕ denotes a non-linear activation function (AF). SoftMax, relu, tanh, sigmoid, and other activation functions are often employed. In the same way, the decoder maps the hidden representation as follows [143]:

$$d = \phi(W' * i + b') \tag{2}$$

In addition to traditional auto-encoders and sparse auto-encoders, there are a few altered variations available, such as denoising AE, contractive AE, and variational AE. Vincent et al. have suggested that a denoising AE can be used as training criteria to learn and extract the essential features, which can yield an efficient high-level feature from the input [151]. An explicit regularizer is introduced to the objective function of a contractive AE, forcing the model to grasp an encoding that is resilient to minor input value fluctuations. The variational auto-encoder is a type of generative model that is classified as an auto-encoder because of its architectural similarity to the basic auto-encoders [152]. The deep auto-encoders offer many advantages, like [143]:

- Deep AE facilitates the reduction in the computational power required for the representation of some functions.
- Deep AE facilitates the reduction in the computational training data required for learning some functions.

AEs are trainable in an unsupervised way. The stacked denoising AE (SDA) can offer an efficient pre-training solution. The model is trained by instantiating the weights of a DNN. When the SDA has been trained layer-by-layer, the auto-encoders' parameters can be used to initialize all of the DNN's layers. Then, on the labelled training data, supervised fine-tuning is used to reduce the prediction error. To map the output of the final layers to the targets, a SoftMax layer is typically placed on top of the AE-based structure; Figure 10 illustrates this step. The pretraining method using SDA improves the convergence speed of the DNN models compared to the random weight initialization. Training of the DNN includes issues like vanishing/exploding gradient problems. This is due to the commonly used nonlinear activation functions (tanh or sigmoid). As a result, auto-encoder-provided unsupervised training is valuable and effective.

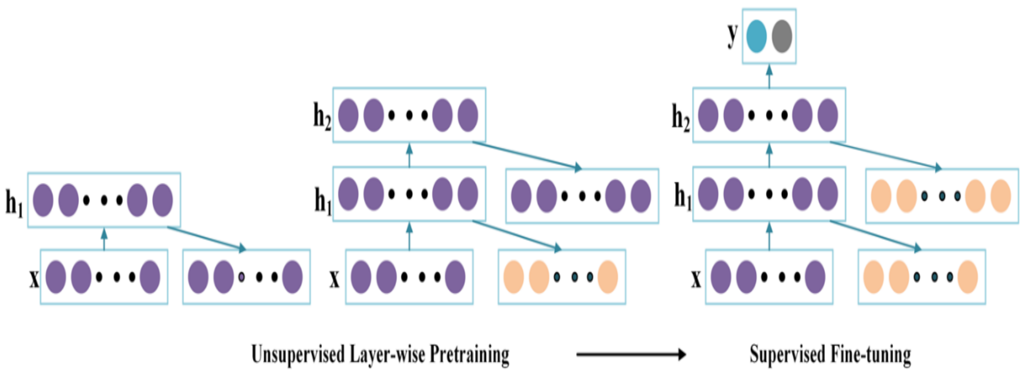

Figure 10. Framework for unsupervised pre-training and supervised fine-tuning of stacked denoising auto-encoder based DNN model.

4.3. Convolutional Neural Network

The visual cortex of the human brain is the inspiration for the convolutional neural network (CNN), which was first proposed by LeCun [153] for image processing. The CNN has found its application in multiple domains, including image segregation, speech recognition, object recognition, and many more. The CNNs are analogous to traditional ANNs as they also consist of neurons that self-optimize through learning. Like ANN, CNN is also composed of multiple layers, including an input layer, hidden layer, and output layer. In the CNN, the layers performing convolution operations are the hidden layers. The CNN is mainly composed of the convolutional layers (CLs), pooling layers (PLs), and fully-connected layers (FCLs) [126]. Figure 11 gives a simple architecture of CNN. The details of these layers are given below:

- CL: This is the core component of a CNN. The majority of the computation occurs in this block only. The input to this layer is the tensor with shape (number of images) × (image height) × (image width) × (input channels). The name convolution comes from the mathematical operation, termed 'convolution', in this layer. Convolution is a linear operation in a CNN that performs a weight multiplication with the input. The CNNs have traditionally been designed for 2-D inputs, with multiplication occurring between a 2-D array of input data and a 2-D array of weights, also known as a kernel or filter. The size of the kernel is a fraction of the input data. Between the filter-sized input matrix and the filter, the dot product is utilized, which is then summed to provide a single value. The tiny-sized filter allows the input array to multiply the same filter (set of weights) several times at various points on the input. The filter is convoluted all over the input data's portion/segment/patch. This is conducted left-to-right and top-to-bottom. The multiplication of the filter and input yields a single value. The input filtering is characterized as a 2-D array of output values obtained by repeatedly applying the filter to the input array. Consequently, a 2-D array obtained through this operation is referred to as a "feature map". The values in the feature map are passed through a non-linearity, such as a Rectified Linear Unit (ReLu), once it has been developed [143]. It can be explained mathematically in the following way:

$$\mathrm{G}[\mathrm{m},\mathrm{n}] = (\mathrm{f} * \mathrm{h})[\mathrm{m},\mathrm{n}] = \sum_{\mathrm{j}}\sum_{\mathrm{k}}\mathrm{h}[\mathrm{j},\mathrm{k}]\cdot\mathrm{f}[\mathrm{m}-\mathrm{j},\mathrm{n}-\mathrm{k}] \tag{3}$$

  where f denotes the input array, h denotes the kernel, j and k denote the input matrix size, and m and n represent the row and column indices of the resultant matrix.
- PL: The PL performs the down-sampling operation, typically applied after a convolution layer. This helps in achieving spatial invariance. It prevents overfitting by aggressively lowering the spatial dimension of the network's representation to decrease the quantum of the parameters and calculations. As it computes a constant input function, it introduces no parameters. In general, max and average pooling are often used in the analysis. Each pooling operation in the max pooling scheme selects the current view's maximum value. Similarly, each pooling action in average pooling averages the current view's value.
- FCL: Similar to conventional neural networks, the FCL neurons are fully connected to the preceding layer. Consequently, a matrix multiplication followed by a bias offset can be utilized to calculate their activations.

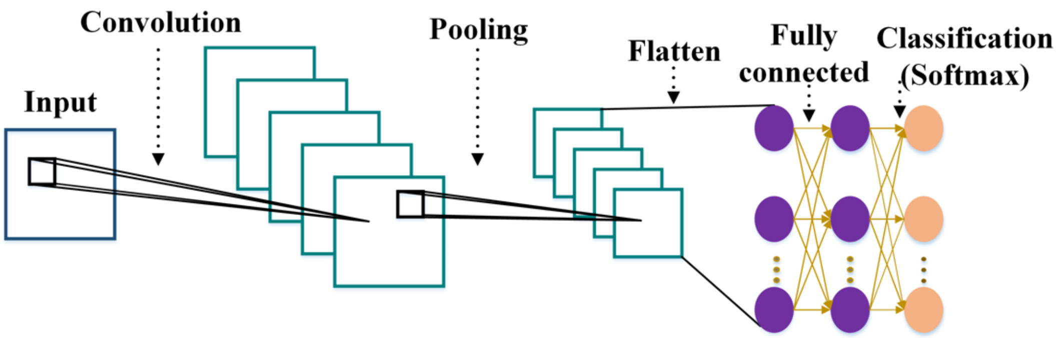

Figure 11. Framework of CNN model.

4.4. Recurrent Neural Network

A recurrent neural network (RNN) is an AI algorithm that performs efficiently on the time series or sequential data. A RNN is well-suited for a variety of problems, like language transformation, natural language processing, voice recognition, and picture captioning. It has found applications in famous applications, like Siri and Google translator. The RNNs, like feedforward and CNNs, learn from training input. They are distinguished by their "memory", which enables them to alter the present action and output by employing information from previous inputs. The RNN differs from other DL algorithms in that its output is reliant on the preceding elements of the input sequence. RNNs use the backpropagation through time (BPTT) technique, which is fundamentally different from traditional backpropagation because it is tailored to sequence data in order to determine the gradients. Traditional backpropagation employs the same concepts as BPTT, wherein the model self-trains by computing the errors from its output to its input layer. These computations allow for the precise modification and adjustment of the model's parameters. BPTT varies from conventional techniques in that errors are accumulated at each time step, whereas feedforward networks do not need to accumulate total errors [143]. This is because feedforward networks do not share parameters between layers. Figure 12 shows the basic structure of a RNN.

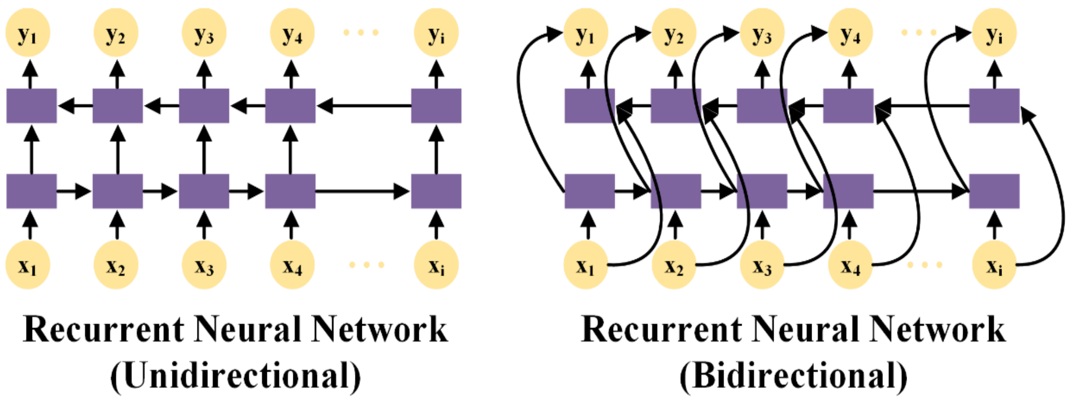

Figure 12. Framework of RNN model.

5. Deep Learning for the PHM of Rotating Machinery of Industrial Robots

This section discusses the existing deep learning framework-based PHM of the rotating machinery of IRs. The existing deep learning frameworks for PHM are discussed in the following sections.

5.1. Deep Belief Network for PHM

The DBN is among the most popular algorithms in the domain of PHM. It is among the first models employed for the PHM strategy. Dash et al. [154] (p. 20) developed a DBN-based probabilistic generative model to detect robotic manipulator failure. Failures have been identified at every position and instance of robotic manipulators using the DBN-based model. Elsewhere, Chen and Li proposed a DBN-based model that utilizes the features extracted with the help of auto-encoders from the vibration data for bearing FDT [155]. Ren et al. propounded a FDT methodology using DBN models [156]. The DBN model was trained on historical datasets and applied to real-time measurement data to generate outputs. These outputs and measurements were used to extract residuals, and based on the adaptive threshold for the residuals, faults were detected in the complex system. Xing et al. proposed an invariant DBN model for gear-FDG with the help of raw vibration data [157]. The propounded fault diagnosis model learned the distributed-invariant features directly, utilizing the raw vibration data, and performed the FDG. Jiao and Zheng proposed a combination of the DBN-based model and wavelet transformation of the vibration signals for the fault diagnosis of industrial robots [158]. The vibration signal was denoised, decomposed, and reconstructed using the wavelet transformation. The normalized eigenvector was developed and used to input the DBN-based model. Elsewhere, Ji et al. proposed a methodology for the FDG of the reducers of industrial robots with the help of the deep-level probability-directed graph DBN model [159]. Shao et al. proposed a FDG technique using a DBN-based model for the motors used in the manufacturing process [160]. The DBN model comprises the stacked RBMs (as shown in Figure 13) and is trained with the help of a layer-by-layer pre-training method. This model assesses the motor's health by automatically learning aspects from the sensor data. Most of these methods still require hand-crafted features and efficient signal processing techniques. These dependencies restrict the model's performance. The training of DBN-based models is cumbersome, and tracking the loss function is challenging. It also limits the provision of end-to-end learning solutions for fault diagnosis and prognosis.

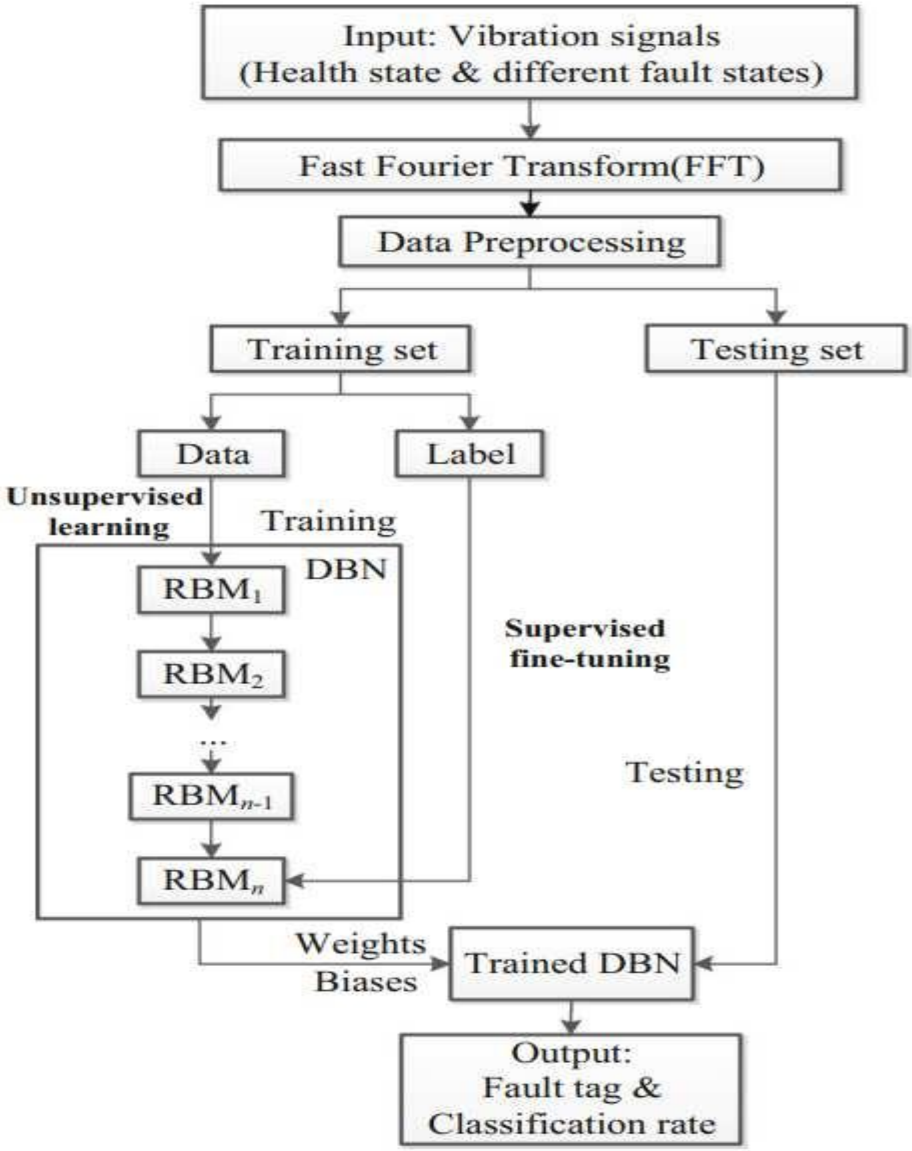

Figure 13. DBN-based framework for fault diagnosis of the induction motor used in manufacturing process [160].

5.2. Deep Boltzmann Machine for PHM

The Deep Boltzmann machines are powerful DL models that conduct the interpretation and training process in cooperation with the bottom–up and top–down directions.

This helps improve the representation of the input features. Despite being a powerful deep learning model, limited works are related to the DBM-based PHM strategies. Elsewhere, Hu et al. proposed a FDG approach for industrial fault diagnosis that included the DBM and multi-grained scanning forest ensemble [161]. Deng et al. proposed a bearing FDG employing the DBM-based FDG paradigm [162]. The time and frequency domain features were retrieved and fed into the DBM-based model as input. Li et al. used the Gaussian–Bernoulli DBM for high-level feature development using the vibration data in three modalities [163]. Figure 14 illustrates the design of the purported framework, and three distinct modalities were fed to the Gaussian–Bernoulli DBM for gearbox FDG. A SVM classifier was utilized to fuse the representative features and perform the fault classification. This approach has been verified on both spur and helical gearboxes. Wang et al. purported the Gaussian–Bernoulli DBM for compressor health management in smart manufacturing [164]. The DBM Gaussian neurons were used to pre-process the vibration signals for health management, and the created model was able to infer the complicated features from the input sequence. Hyperparameter optimization was carried out using the Particle Swarm Optimizer algorithm. Also, a tailored Liu–Storey conjugate gradient algorithm was used to improve the convergence rate. The application of DBM-based models requires intensive computation, and conducting a weight update is challenging. However, DBM-based models also have advantages, like efficient learning of complex representations and good uncertainty propagation. Mitrevski and Ploger [165] have proposed robot fault detection and diagnosis using the DBM-based model.

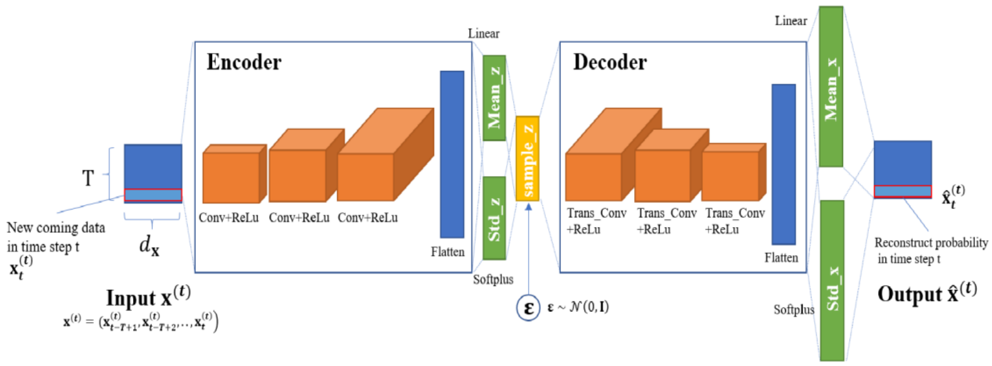

Figure 14. Gaussian–Bernoulli DBM framework for fault diagnosis [166].

5.3. Auto-Encoder for PHM

Auto-encoders are very popular in the PHM field, and have been extensively used by researchers for FDG and FP. Hong et al. presented a FDG approach for multi-joint industrial robots using a deep sparse auto-encoder model using an attitude dataset [167]. A test rig was used to help create the dataset, and analysis was performed on the results. Reference [166] proposed a fault diagnosis approach for the industrial robot using the sliding-window convolutional variational auto-encoder-based model and multivariate time series data. Figure 15 shows the framework of this approach, where the input data is $x^{(t)}$ with a time-step of t, and the model output is the reconstruction probabilities of each point in the sliding window. Xiao et al. propounded a denoising AE-based model using acoustic signals for fault diagnosis [168]. Elsewhere, Yun et al. proposed a fault diagnosis approach for the robot arm using the stacked auto-encoder (SAE)-based FDT framework [169]. Two stethoscopes were used for sound data acquisition, and feature extraction was performed using the STFT spectrogram. The auto-encoder model was trained and tested using these features. Sun et al. developed a FDT approach for induction motors using a sparse AE for the feature learning from the vibration signals [170]. Partial corruption was added using the denoising coding and fed into the SAE-based model for feature learning. These features have been fed to the neural network classifier for fault identification and classification. In another work, Li et al. proposed an intelligent FDG using a fully connected AE for bearing FDG [171]. The proposed model imposed a lifespan sparseness on the encoded features, and the soft polling method was applied to boost the accuracy and stability. Also, a dataset was developed by adding Gaussian noise, and the performance was evaluated to validate the performance in a noisy environment. Sohaib et al. developed an approach for the FDG approach for rotary machine bearings with the SAE model [172]. The complex envelope spectrum aided in making the frequency component more distinct in the signal and facilitated efficient feature extraction from the given input signals. Also, automatic feature extraction helped in tackling the problems associated with manual feature extraction and selection. AE-based approaches require a dedicated classifier for fault diagnosis. These approaches require a high computational cost, and selecting the specific features for the fault diagnosis is challenging. However, AE-based techniques also offer advantages, like flexible architecture, dimensionality reduction, and no requirement for labeled data.

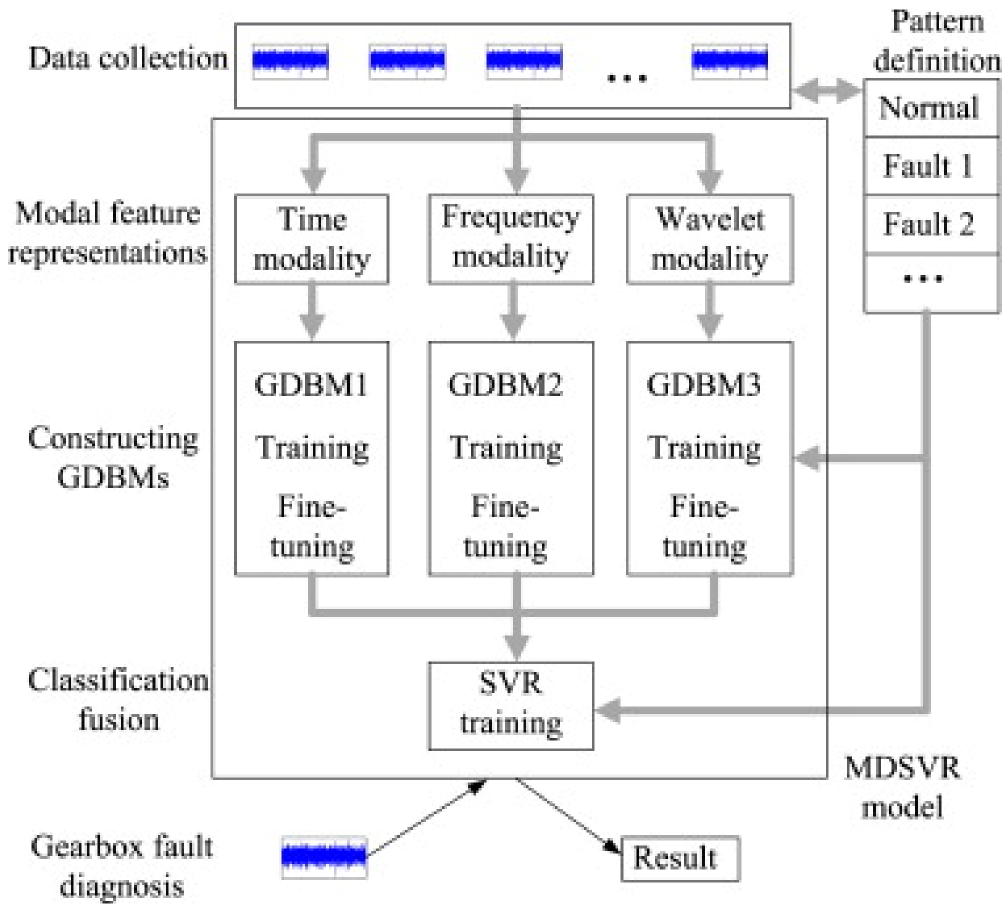

Figure 15. Convolutional variational auto-encoder framework for fault diagnosis in industrial robots [163].

5.4. Convolutional Neural Network for PHM

The CNN is efficiently used in PHM due to its excellent automatic feature learning capabilities and segregation capabilities. Chen et al. developed a CNN-based FDG model for the heavy-duty industrial robot system [173]. A FDT model was designed with a series combination of spectrum calculation fault diagnosis networks. Elsewhere, Kim et al. suggested a FDG methodology for industrial robot servo systems utilizing multiple sensor signals as an input to the one-dimensional CNN model [174]. Figure 16 shows the framework, and it is comprised of one plain convolutional block, three stacked residual blocks with a skip connection, a global average pooling (GAP) layer, and two fully connected layers. Also, the model was validated on the Case Western Reserve University (CWRU) data and the IMS bearing database of the University of Cincinnati. Yang et al. developed a FDT technique for the rotating vector reducer for industrial robots using a CNN-based model [175]. Meanwhile, Ma et al. [176] suggested a FDT methodology for the industrial robot by employing the one-dimensional CNN and data improvisation through random sampling and mix-up data augmentation. The dataset includes the torque, speed, position, and current data of the robot. Li et al. proposed a multi-axis IR FDG approach using the multi-label one-dimensional CNN [177]. Elsewhere, Liu at al. proposed a dilated CNN model for the cross-axial industrial robotics FDG [178]. The sliding window and key feature extraction method pre-processed the input data. These data are fed to the dilated CNN model for feature mining, and the self-attention network is used for its feature attention capability. Lu et al. propounded a dual-module attentive CNN for industrial robot fault diagnosis [179]. Two parallel CNN models with different attentions were capitalized for feature learning, and the features were fused for efficient fault diagnosis. In another work, Janssens et al. proposed a CNN-based technique for different types of bearing faults and rotor imbalances [180]. The CNN-based FDT approach helped eradicate the need for manually engineered features like the ball pass frequencies of the raceway, kurtosis, RMS, and variance. It demonstrated that a CNN-based feature learning system outperforms classic feature-engineered techniques. Plakias et al. developed an attentive dense CNN fault diagnosis technique for bearings [181]. The attentive deep CNN model considers the temporal coherence of the data, which helps to improve the feature learning. Cheng et al. developed a FDT technique for rotating machinery using the CNN model and continuous wavelet transformation [182]. The proposed model used a local binary convolutional layer for faster training and to avoid overfitting issue. The model was tested on the bearing faults and gearbox compound FDG. Guo et al. developed a CNN-based FDT approach for rotating machinery employing the continuous wavelet transformation of vibration signals [183]. Also, the model was tested on different pieces of rotor equipment for validation of the proposed methodology. Liang et al. purported an intelligent FDT approach for rotating machinery using the combination of the wavelet transformation, generative adversarial nets, and CNN [184]. The wavelet transformation was used to obtain the time–frequency image characteristics from the one-dimensional raw signals. Generative adversarial nets were employed to develop the training images, and the CNN model was used for FDT. Li et al. propounded a FDG approach for rotating machinery with the amalgamation of the DBN and one-dimensional CNN [185]. The DBN was created by combining the three RBMs for high-dimensional data feature abstraction and dimensionality reduction. For FDG, these low-dimensional features are loaded into a one-dimensional CNN model. To evaluate the effectiveness of the model, the proposed approach was applied to two experimental datasets. These CNN-based methods facilitate efficient feature extraction and safeguard spatial information. Also, the performance of these models relies on parameter initialization and lacks global feature extraction capabilities.

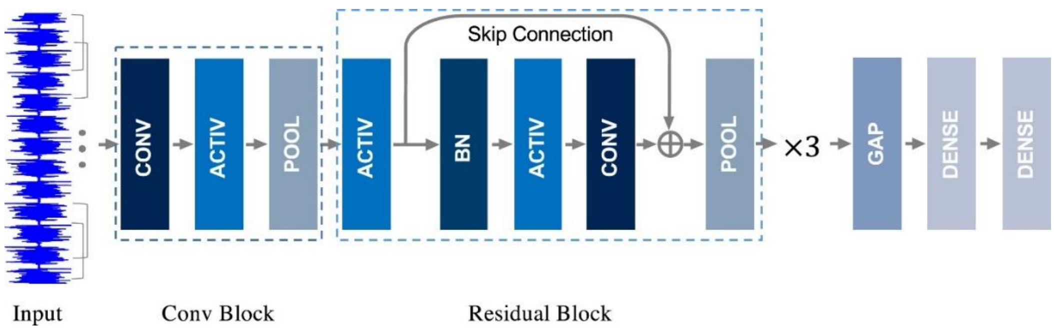

Figure 16. Residual CNN framework for fault diagnosis in industrial robots [174].

5.5. Recurrent Neural Network for PHM

The RNN provides the inherent advantage of retaining the temporal information of the time–series data. This property of the RNN is advantageous for PHM as it retains the information throughout the training. The RNN can recall temporal dependencies and understand the failure's dynamic comportment. Vanilla RNNs (basic RNNs), on the other hand, are unable to learn long-term temporal dependencies due to the vanishing/exploding gradient problem. Gradient clipping is a technique that uses a threshold value to restrict the magnitude of a gradient. Various gating strategies have been proposed to deal with the vanishing gradient. Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU) are the two most well-known RNN versions for dealing with these problems. For example, An et al. developed a bearing FDT method with the help of an RNN-based FDT model [186]. Figure 17 shows a framework of the propounded model and training strategy. The input network extends the dimensions of the input, LSTM cells have been used for the recurrent framework, and hidden inputs are used as an input to the two-layer network. Also, the physical interpretation of the network learning was given with the help of the maximum mean discrepancy and t-SNE [187]. Zhang et al. developed an RNN-based FDT model for rotating machinery [188]. GRU is used to extract temporal information from the time–series data and to learn the relevant attributes from the produced images. Finally, MLP is used to implement FDT. Liu et al. propounded a FDT method for the rolling element bearing using the recurrent neural network [189]. The reconstruction error between the output data and the following period data was utilized to detect and categorize different fault kinds after the vibration signals were employed as an input to the GRU-based denoising AE-based model. Elsewhere, Jaing et al. developed a bearing FDT method using the deep RNN model [190]. The frequency signals are used to construct the input data without the manual feature development, and the adopted DL strategy was applied for the training process. Qiao et al. suggested a FDT model through the combination of the CNN and LSTM for bearing fault diagnosis [191]. The spatial sequence characteristics were extracted from the input signal by the model's convolution and LSTM layers, and the final classification was performed by the dense layers. Also, the model was substantiated on the public domain data sets of the CWRU. Meanwhile, Oh et al. proposed a bearing fault diagnosis of the rotating machinery using the combination of a denoising AE and a multi-scale convolution RNN [192]. The denoising AE is used to pre-process the data, and the multi-scale convolution RNN is applied to categorize the bearing defects. Li et al. [193] proposed a mobile robot bearing FD using the DWT and LSTM models. The vibration signals were decomposed into six frequency bands for the bearing FD task. Zhi et al. [194] developed a FD model for harmonic reducers, which is a key component of IRs. A combination of the CNN and LSTM has been employed for the FD using the wavelet regional correlation threshold denoising algorithm. Wang et al. [195] proposed a FD method for the motor drive system of IRs, where the CNN has been used for the feature extraction and LSTM for the prediction of the system's health. RNN-based approaches are among the most efficient methods for RUL estimation. The RNN is well suited to time–series data and sequential data. The conventional RNN structure often suffers from the vanishing gradient problem. However, a modified version of the RNN, like LSTM, helps solve the vanishing gradient problem. The training time of these models requires a long duration, and the structures are often complex. Generally, these models do not support parallel computing. Overall, a few DL-based PHM strategies have been encapsulated in Table 2.

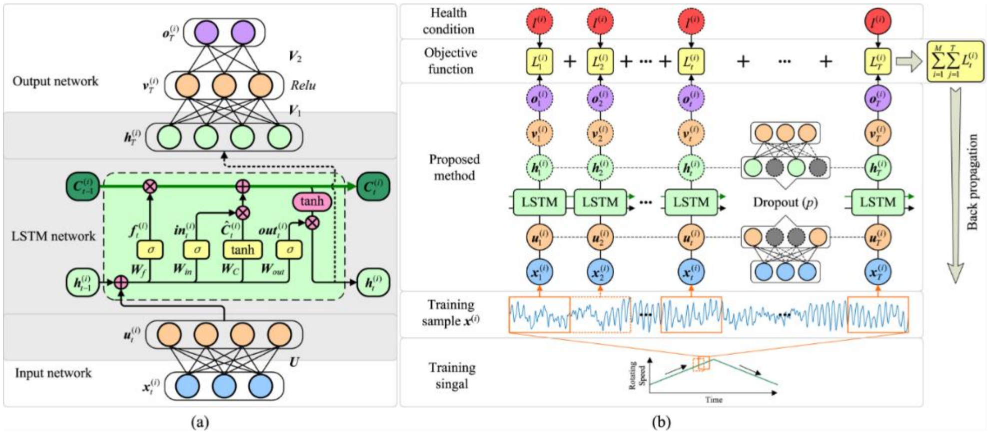

Figure 17. Illustration of (**a**) RNN framework for fault diagnosis in bearings, and (**b**) training methodology [186].

Table 2. PHM Methods with DL-based strategies.

PHM Methods Based on Deep Learning Methods					
	Reference	**Dataset**	**Datatype**	**Model**	**Accuracy**
Fault diagnosis	Sohaib et al. [196]	CWRU bearing data	Vibration	SAE-DNN	99.5%
	Xu et al. [197]	CWRU and Xi'an Jiaotong University (XJTU-SY) bearing data	Vibration	Wavelet Transform-CNN	99.4%
	Hoang and Kang et al. [198]	KAT bearing dataset by Paderborn University, Germany	Current	CNN	99.47%
	Li et al. [199]	CWRU	Vibration	CNN-LSTM	99.74%
	Chen et al. [200]	Gear Box	Vibration	DBM	99.94%
		Gear Box	Vibration	SAE	99.55%
		Gear Box	Vibration	DBN	98.73%
	Chen et al. [201]	CWRU	Vibration	Cyclic Spectral Coherence-CNN	98.93%
	Verstraete et al. [202]	CWRU	Vibration	2D-CNN	99.9%
	Liu et al. [203]	Test setup with motor and faulty bearings	Vibration	STFT-SAE	97.84%
	Shi et al. [204]	UNSW planetary test rig	Vibration	Bidirectional-convolutional LSTM	95.83%
	Ravikumar et al. [205]	Test setup with IC engine gearbox	Vibration	Stacked LSTM	94.33%
	Chen et al. [173]	IR system	Current	Improved CNN	99.59%
	Long et al. [206]	IR system	Attitude data	SAE-SVM	96.74%
Fault prognosis	Kamat et al. [207]	PRONOSTIA bearings dataset	Vibration	AE-LSTM	90%
	Wang et al. [208]	PRONOSTIA bearings dataset	Vibration	Spatiotemporal-3DCNN	98.25%
	Ding et al. [209]	NASA IMS dataset	Vibration	Deep CNN	0.0052 (RMSE)
	Li et al. [210]	Test setup with milling machine	Current	LSTM	0.0950 (RMSE)
	Qin et al. [211]	Test setup with gear	Vibration	Macroscopic-microscopic attention LSTM	0.142 (NRMSE)

6. Discussion, Challenges and Future Aspects of PHM

Prognostics and Health Management using DL has proved to be a fast-growing field, with a lot of studies being conducted on it worldwide. It has brought significant improvement and has proven to be a key technology to improve system reliability. A good PHM strategy provides strong economic benefits, minimum downtime, and maximum productivity. The main objective that the PHM system must accomplish, depending on the applications, poses challenges for their requirements. The most obvious ones are accuracy and precision. It needs to be quantified with the set of performance indicators and evaluated against their decision-making. In certain instances, incredibly superior accuracy and precision are essential to making a confident decision. Such a decision may include stopping a system in case of a warning of FDT, replacement of the component upon FDG, and delaying or predicting planned maintenance based on the RUL estimations. High levels of precision and accuracy might not be necessary in some circumstances, and they might conflict with other goals. In some circumstances, the decision-making in the crucial applications of the IR systems depends on the transparency, explicability, and interpretability of the PHM models. PHM solutions offer a practical and intelligent approach to condition-based maintenance and preventive maintenance, but they also have some security flaws. The PHM's technological foundation is made up of numerous devices and different communication protocols. There is always an issue of data integrity, data privacy, and substantiation. These issues are required to be properly addressed to ensure the robustness and efficiency of the PHM solution. Inaccurate PHM models, data availability issues, limited knowledge of the machine's current state, randomness in the machine's future usage profile, and unreliable sensor data values are some issues that arise when using PHM in real-world applications. Also, it is crucial to understand the feasibility of the different sensor signals and signal analysis used for the PHM of IRs. This has been included in Table 3 to demonstrate the signals used, along with the signal processing tools and their pertinency.

Moreover, the characteristics of the rotating machinery in IRs and traditional rotating machinery can differ in a number of ways. Some of the differences are given below:

- Dimension and scale: Compared to conventional rotating machinery, IRs often possess smaller, more compact rotating machinery. Traditional rotating machinery can be substantially larger in size and have higher power ratings, such as large-scale industrial machinery or power generation turbines. Industrial robots, on the other hand, need smaller motors and systems to carry out their specialized tasks quickly and accurately.
- Operating Speed and Precision: Due to the high-speed, precise tasks that IRs are built for, their rotating machinery must operate with outstanding speed and control precision. These robots frequently carry out actions that require the quick acceleration and deceleration of the rotating components.
- Duty Cycle and Constant Operation: IRs frequently work in cycles or sequences, carrying out particular tasks intermittently with brief spikes in activity. Their rotating equipment must be able to resist repeated start-stop cycles and adjust to different load scenarios. Turbines used in power generation or other types of conventional rotating machinery frequently run continuously for long periods of time without undergoing repeated start-stop cycles.
- Maintenance and Serviceability: The frequent and demanding actions of industrial robots necessitate frequent maintenance and service. These robots' rotating machinery needs to be built with access points for inspections, lubrication, and component replacement in mind. Depending on their individual applications, traditional rotating machinery may have distinct maintenance needs that necessitate more involved maintenance processes and extended downtime intervals.

Table 3. Signals and signals processing applicable to the PHM of IRs with their applicability.

Article	Signal	Signal Processing	Feasibility
[212]	Encoder	Singular spectrum analysis (SSA) Hilbert transform (HT) Empirical mode decomposition HT (EMDHT)	Easy availability of signal Simple to apply
[213]	Vibration	Wavelet transform (WT)	Signal easily accessible with sensors Highly feasible
[214]	Vibration	WT	Signal easily accessible with sensors Easy to apply
[215]	Acoustic Emission (AE)	WT	Easily accessible Signal interpretation is cumbersome High background noise Generalization is difficult
[216]	Current	WT	Non-intrusive approach Highly feasible Multi-resolution analysis
[217]	Vibration	Discrete WT	Multi-resolution analysis Localization of Fault Signatures Noise Suppression
[218]	Vibration	Continuous WT	Time-Frequency Localization Continuous Frequency Coverage Adaptability to Signal Variability Improved Accuracy in Transient Detection
[48]	Current	Statistical analysis	Simple applicability Efficient feature development Feature selection needs expertise
[56]	Current	Discrete WT	Non-invasive approach Early fault detection Real-time monitoring
[219]	Vibration	Discrete WT	Multi-resolution analysis Localization of Fault Signatures Efficient feature extraction
[69]	Vibration	Short-time Fourier Transform (STFT) Wavelet decomposition	Time-Frequency localization Simplicity and Computational Efficiency
[220]	Current	STFT Wavelet packet decomposition (WPD)	Non-invasive approach Early fault detection Wide applicability Cost effective
[221]	Vibration	Discrete WT	Multi-resolution analysis Localization of Fault Signatures Interpretability
[222]	Vibration	STFT	Efficient Frequency Resolution Interpretability Fast computation
[169]	Vibration	STFT	Efficient Frequency Resolution Noise suppression Easy to implement

Deep learning has brought significant improvements to the PHM approaches by learning complex representations of the data. The capability of DL frameworks to automatically extract and learn features provides an edge over the conventional machine learning models.

Despite multiple advantages, DL-based PHM still has a long way to go. The DL algorithms, like RBM, DBN, DBM, AE, CNN, and RNN, have brought significant improvements to the PHM strategies. Table 4 compares the different DL algorithms. The algorithms are chosen based on the type and availability of the data. Only a small amount of research work is available for the PHM strategies with DL applications for industrial robots. However, many PHM strategies with DL applications for rotating machinery are available. In recent years, researchers worldwide have used DL algorithms for FDT, FDG, and FP. The DL algorithms have challenges associated with them too. Many improvements are needed for industrial acceptance of the PHM strategies with DL algorithms, and questions need to be answered. Here are some of the important aspects of the DL-based PHM strategies:

- Data insufficiency: In the real-time environment, the unavailability of a large amount of data is the major hurdle that restricts the application of PHM strategies with DL algorithms. It is well-known that DL algorithms require a substantial volume of data, and that the availability of large volumes of data is not feasible. Some of the established DL frameworks, like VGG16, VGG19, ResNe-50, and InceptionResNet-v2, have used millions of images for training [223–225]. However, in a real-time context, such a massive amount of data is not feasible. Researchers have used techniques like data augmentation to increase the training samples in the training datasets by developing synthetic data. Basic data augmentation methods like window cropping, wrapping, and flipping can be applied to time-series data to create a variety of data structures [226]. Generic algorithms are also frequently used to create new data that is similar to the original data. In some cases, new GMs are used to construct the linked time-series data, keeping the temporal dependency of the original data. Researchers have also looked into the concept of transfer learning to address the data insufficiency problem for FDG and FP [227,228]. Many new innovations are coming, and they will aid in improving the PHM strategies with DL applications.
- Data quality: The performance of any artificial intelligence-based model relies heavily on the data quality. The success and efficiency of DL-based PHM strategies depend heavily on data quality. The availability of cloud computing, the industrial internet of things, and intelligent sensors have aided in collecting an enormous amount of data. However, the growing volume of data brings included noises and disturbances. Also, the ambient and operating conditions affect the quality of the data. In industrial data, there are the problems of data duplicity, unlabeled data, imbalanced data, and many more. These concerns have not been addressed thoroughly in most of the available research works. Also, the majority of solutions focus on a minimally imbalanced scenario, ignoring the problems associated with the substantially under-represented instances, which are common in real-world industrial workplaces. Furthermore, real-world data is generally unstructured, multi-modal, and diverse, making the model far more challenging. More attention is required in the future for developing a generic deep model that can work on diverse data without sacrificing the training efficiency.
- Data pre-processing: This is among the most crucial components of an AI-based model's effective performance. It is crucial for both machine learning and DL models. The model's success is mainly reliant on the condition of the input data. Pre-processing includes data normalization, the removal of data duplicity, and standardization. It also includes tackling incomplete data problems, outliers and missing values, and labeling data. In certain cases, signal processing tools, like FFT, STFT, wavelet transformation, and Hilbert–Huang transformation, are also used to process the input signals. In the future, research will be required to build a standardized approach for pre-processing data prior to their input to the AI-based model.
- Model selection and explainability: The appropriate DL framework selection is one of the vital steps for the development of an efficient PHM strategy. There are many DL algorithms, and choosing the appropriate model based on the available data is critical. Most of the available PHM strategies have been based on the models that require handcrafted features. These models are prone to errors and lack generalization

capabilities. The application of DL algorithms has helped to resolve the problems associated with the handcrafted features. However, setting up the hyper-parameters of the DL-based model is itself a big challenge. There are only a few papers that deal with setting up the DL models and their hyper-parameters' optimization. There is a need to develop a strategy that would allow the autonomous optimization of models and their hyper-parameters as per the given input data. This can be investigated in the future. Despite the good performance of DL models for PHM strategies, the acceptance of such an approach has a lot of roadblocks. The DL models are like black box models and lack interpretability. The decision-making part of PHM is heavily dependent on the DL models. However, only a few papers have dealt with the interpretability and explainability of DL models for PHM [126]. There is a need to develop a PHM strategy with DL applications, but with transparency, interpretability, and explainability.

Table 4. Pros and cons of DL-based PHM.

Algorithms	Pros	Cons
DBN [78–80]	Global feature extraction possible	Training is cumbersome
	Supports dimensionality reduction	Tracking loss function is difficult
	Can work on less data	Parameters optimization is difficult
DBM [161,163]	Can learn internal representation	Weight update is difficult
	Robust to ambiguous inputs	Slow training
CNN [36,96,98,115]	Excellent feature extraction properties	Cannot obtain global features
	Supports multiple dimension data	Cannot interpret time dimension information
Auto-encoder [166,168]	Unsupervised learning	Requirement of dedicated classifier for fault diagnosis
	No label data requirement	High data requirement
	Supports dimensionality reduction	Difficult to determine the importance of data
	Supports flexible framework	Selecting specific features not possible
	Availability of multiple forms	No interpretability
RNN [109,111–113]	Performs well on sequence problem	Slow training speed
	Capitalizes the time dimension of input data	No parallel computing
	Supports unlimited input length data	Problem of vanishing gradient

The application of PHM to IRs offers several challenges. It can be more difficult compared to applying it to the other rotating machines or components. Some of the challenges pertaining to the PHM of IRs are listed below:

- Industrial Robots' Complexity: IRs are sophisticated systems with numerous vulnerable parts. This makes it challenging to gather and examine the data that might be used to anticipate problems.
- Dynamic Operating Conditions: IRs are frequently utilized in a range of settings with various operating circumstances. Because of this, creating PHM models that can correctly forecast failures under all operating circumstances can be challenging.
- Lack of Sufficient Training Data: For the development of reliable FDT and prediction models, it can be difficult to obtain enough labelled training data. In IR systems, labelled data collection can be time-consuming and expensive for different fault scenarios. Furthermore, gathering data for uncommon or catastrophic failure situations might be very difficult.

- Adaptability and Generalization: Systems for IRs might differ greatly in terms of their models, configurations, and working environments. PHM systems must be flexible and able to be generalized to various robot kinds and settings. To achieve dependable and scalable PHM, it is challenging to create models and algorithms that can adjust to differences in robot systems, including changes in the load, operational conditions, or task requirements.

Deep learning algorithms in PHM systems built on the DL framework have a number of possibilities for development. The following are some ideas for potential enhancements:

- Model Architectures: The performance can be enhanced by investigating and creating new model architectures designed specifically for PHM workloads. This entails creating deep neural networks with specialized layers, such as recurrent or attention processes that can efficiently capture temporal dependencies and persistent patterns in the sensor input. In order to capture complicated interactions in multi-modal or graph-structured data, architectural innovations like transformer models or graph neural networks can also be researched.
- Uncertainty Quantification: DL models often lack the ability to provide reliable uncertainty estimates, which is crucial for decision-making in PHM systems. Model uncertainty can be better understood and characterized by including uncertainty quantification approaches, like Bayesian deep learning or Monte Carlo dropout. As a result, it will be easier to spot circumstances when the model's predictions may not be as accurate and make the appropriate decisions.
- Robustness to Adversarial Attacks: Small changes in the input data can cause inaccurate predictions or misclassification in deep learning models, making them vulnerable to adversarial attacks. To increase the resilience of deep learning models in PHM systems, adversarial robustness strategies might be investigated, such as adversarial training or input regularization. With the help of these methods, models should be more resistant to adversarial examples and perform consistently, even when there are subtle attacks or data abnormalities present.

PHM is a promising technology that has the potential to increase the dependability and uptime of industrial robots despite these difficulties. The technology is anticipated to become more accessible and less expensive as it advances. As a result, PHM will be a more appealing option for more producers. IR prognostics and health management is a developing field with a number of promising future developments. Future PHM capabilities for industrial robots include the following:

- Real-time Monitoring and Adaptive Control: Future PHM systems for IRs will place a strong emphasis on real-time monitoring and adaptive control techniques. Continuous assessment of the robot's health is made possible via real-time monitoring, allowing for the quick identification and remediation of any potential problems. In response to recognized flaws or degradation, adaptive control approaches can modify the robot's operational settings or control algorithms, improving the performance and lowering the likelihood of failure.
- Human–Robot Collaboration and Safety: Collaboration between people and robots in shared workspaces will rise in the future of industrial robotics, as will safety concerns. The safety and wellbeing of human operators will be greatly enhanced by PHM systems. PHM systems can initiate safety routines, such as reducing the robot speed, changing the motion trajectories, or shutting down the robot in critical conditions, by monitoring the robot's health and identifying probable defects. For the creation of secure and effective human–robot collaborative environments, this PHM feature will be crucial.
- Predictive Maintenance and Spare Parts Management: Predictive maintenance and improved spare parts management are the future directions for PHM systems for IRs. These systems can calculate the remaining useful life of crucial components and schedule maintenance procedures appropriately by utilizing predictive models and

real-time monitoring. This method ensures that IR systems operate well by minimizing unplanned downtime, lowering the maintenance costs, and optimizing the spare parts inventory.

- Cloud Computing and Remote Monitoring: The use of cloud computing and remote monitoring tools will make it possible to centrally monitor and analyze a number of IRs spread out across several places. Distributed robot data can be gathered and analyzed via cloud-based PHM platforms, enabling benchmarking, trend-tracking, and comparison analysis. Experts can remotely monitor and help with IR system troubleshooting thanks to remote access and diagnostics, which promote proactive maintenance and support.
- Big Data Analytics and Deep Learning: IR sensor data is becoming more widely available, creating new opportunities for using big data analytics and DL methods. Large amounts of sensor data can be analyzed to find patterns and anomalies that improve issue identification and diagnosis. On the basis of the previous data, DL algorithms can be trained to create predictive models that can foretell errors or performance declines in the future. PHM systems for IRs may be more accurate and reliable when using this data-driven approach.

7. Conclusions

This paper has presented a review of the PHM strategies with different deep learning algorithms. The application of different DL algorithms, like RBM, DBN, DBM, AE, CNN, and RNN, in PHM for the rotating machines of industrial robots has been discussed, along with the brief theoretical aspects of the algorithms. DL algorithms have significantly improved the performance of the PHM strategies. However, the industrial application of such approaches requires further improvement in terms of the high data requirements, required computational power, and model optimization. The optimization of deep models is a difficult task and involves significant improvements to be made. With the availability of an enormous volume of data with intelligent sensors, cloud computing, and IIoT, PHM with DL algorithms for industrial robots and its rotating machinery is undergoing a huge boost. PHM for industrial robots has a very bright future. The following possibilities could arise as technology advances:

- More Accurate and Reliable Predictions: DL algorithms will grow more potent and sophisticated, enabling us to make forecasts regarding the health of industrial robots that are more dependable and accurate. Less unplanned outages and downtime will result from this, which will increase productivity and save organizations money.
- Improved Decision-making: PHM systems will provide industries with better data concerning the condition of their robots, enabling them to choose more wisely between maintenance and repairs; operations will become more effective and efficient as a result.
- Earlier Detection of Failures: Early failure detection using PHM systems will provide companies more time to take corrective action. Catastrophic failures, which can be expensive and harmful, will be less likely as a result.

Moreover, there is also a need to develop an effective, transparent, interpretable, and explainable PHM approach with DL applications for the rotating machinery of industrial robots.

Author Contributions: Conceptualization, P.K., S.K. and H.S.K.; methodology, P.K.; software, P.K. and S.K.; validation, P.K., S.K. and H.S.K.; formal analysis, P.K.; investigation, P.K.; resources, H.S.K.; data curation, P.K. and S.K.; writing—original draft preparation, P.K. and S.K.; writing—review and editing, P.K., S.K. and H.S.K.; supervision, H.S.K.; funding acquisition, H.S.K. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by a project for Smart Manufacturing Innovation R&D funded the Korea Ministry of SMEs and Startups in 2022 (Project No. RS-2022-00140460).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of the data; in the writing of the manuscript; or in the decision to publish the results.

References

1. Afsari, K.; Gupta, S.; Afkhamiaghda, M.; Lu, Z. Applications of collaborative industrial robots in building construction. In Proceedings of the 54th ASC Annual International Conference Proceedings, Minneapolis, MN, USA, 18–21 April 2018; pp. 472–479.
2. Evjemo, L.D.; Gjerstad, T.; Grøtli, E.I.; Sziebig, G. Trends in smart manufacturing: Role of humans and industrial robots in smart factories. *Curr. Robot. Rep.* **2020**, *1*, 35–41. [CrossRef]
3. Rao, J. *History of Rotating Machinery Dynamics*; Springer Science & Business Media: Berlin/Heidelberg, Germany, 2011; Volume 20.
4. Ruishu, Z.; Chang, Z.; Weigang, Z. The status and development of industrial robots. In Proceedings of the IOP Conference Series: Materials Science and Engineering, Kuala Lumpur, Malaysia, 13–14 August 2018; IOP Publishing: Bristol, UK, 2018; Volume 423, p. 012051.
5. Singh, B.; Sellappan, N.; Kumaradhas, P. Evolution of industrial robots and their applications. *Int. J. Emerg. Technol. Adv. Eng.* **2013**, *3*, 763–768.
6. *Industrial Robotics*; Industrial Press: South Norwalk, CT, USA, 2008; Available online: https://books.industrialpress.com/9780831133580/industrial-robotics (accessed on 29 November 2022).
7. Engelberger, J.F. *Robotics in Practice: Management and Applications of Industrial Robots*; Springer Science & Business Media: Berlin/Heidelberg, Germany, 2012; ISBN 978-1-4684-7120-5.
8. Nof, S.Y. *Handbook of Industrial Robotics*; John Wiley & Sons: Hoboken, NJ, USA, 1999; ISBN 978-0-471-17783-8.
9. Chatterjee, P.; Manikrao Athawale, V.; Chakraborty, S. Selection of industrial robots using compromise ranking and outranking methods. *Robot. Comput.-Integr. Manuf.* **2010**, *26*, 483–489. [CrossRef]
10. Choi, S.; Zhang, G.; Fuhlbrigge, T.; Watson, T.; Tallian, R. Applications and requirements of industrial robots in meat processing. In Proceedings of the 2013 IEEE International Conference on Automation Science and Engineering (CASE), Madison, WI, USA, 17–20 August 2013; pp. 1107–1112.
11. Karabegović, I. The Role of Industrial Robots in the Development of Automotive Industry in China. *Int. J. Eng. Work.* **2016**, *3*, 92–97.
12. Bader, F.; Rahimifard, S. A methodology for the selection of industrial robots in food handling. *Innov. Food Sci. Emerg. Technol.* **2020**, *64*, 102379. [CrossRef]
13. Olsson, T.; Haage, M.; Kihlman, H.; Johansson, R.; Nilsson, K.; Robertsson, A.; Björkman, M.; Isaksson, R.; Ossbahr, G.; Brogårdh, T. Cost-efficient drilling using industrial robots with high-bandwidth force feedback. *Robot. Comput.-Integr. Manuf.* **2010**, *26*, 24–38. [CrossRef]
14. Zhang, G.Q.; Li, X.; Boca, R.; Newkirk, J.; Zhang, B.; Fuhlbrigge, T.A.; Feng, H.K.; Hunt, N.J. Use of Industrial Robots in Additive Manufacturing—A Survey and Feasibility Study. In Proceedings of the ISR/Robotik 2014, 41st International Symposium on Robotics, Munich, Germany, 2–3 June 2014; pp. 1–6.
15. International Federation of Robotics. *Executive Summary World Robotics 2020 Industrial Robots.* 2020.
16. Industrial Robots Market Size, Share & Growth Report. 2029. Available online: https://www.fortunebusinessinsights.com/industry-reports/industrial-robots-market-100360 (accessed on 22 May 2023).
17. Global Industrial Robotics Market Size & Share Report. 2030. Available online: https://www.grandviewresearch.com/industry-analysis/industrial-robotics-market (accessed on 22 May 2023).
18. Jaber, A.A. *Design of an Intelligent Embedded System for Condition Monitoring of an Industrial Robot*; Springer: Berlin/Heidelberg, Germany, 2016.
19. Scott, M.J.; Verhagen, W.J.C.; Bieber, M.T.; Marzocca, P. A Systematic Literature Review of Predictive Maintenance for Defence Fixed-Wing Aircraft Sustainment and Operations. *Sensors* **2022**, *22*, 7070. [CrossRef]
20. Arnaiz, A.; Iung, B.; Adgar, A.; Naks, T.; Tohver, A.; Tommingas, T.; Levrat, E. Information and Communication Technologies Within E-maintenance. In *E-Maintenance*; Holmberg, K., Adgar, A., Arnaiz, A., Jantunen, E., Mascolo, J., Mekid, S., Eds.; Springer: London, UK, 2010; pp. 39–60; ISBN 978-1-84996-205-6.
21. Sheppard, J.W.; Kaufman, M.A.; Wilmer, T.J. IEEE Standards for Prognostics and Health Management. *IEEE Aerosp. Electron. Syst. Mag.* **2009**, *24*, 34–41. [CrossRef]
22. *Condition Monitoring and Diagnostics of Machines. Data Processing, Communication and Presentation: General Guidelines*; BSI British Standards: London, UK, 2003.
23. *Condition Monitoring and Diagnostics of Machines. Data Processing, Communication and Presentation: Data Processing*; BSI British Standards: London, UK, 2015.

24. *Condition Monitoring and Diagnostics of Machines. Data Processing, Communication and Presentation: Communication*; BSI British Standards: London, UK, 2007.
25. Medjaher, K.; Zerhouni, N.; Gouriveau, R. *From Prognostics and Health Systems Management to Predictive Maintenance 1: Monitoring and Prognostics*; John Wiley & Sons: Hoboken, NJ, USA, 2016; ISBN 978-1-119-37102-1.
26. Pecht, M.G.; Kang, M. *Prognostics and Health Management of Electronics: Fundamentals, Machine Learning, and the Internet of Things*; John Wiley & Sons: Hoboken, NJ, USA, 2018; ISBN 978-1-119-51535-7.
27. Kim, N.-H.; An, D.; Choi, J.-H. *Prognostics and Health Management of Engineering Systems*; Springer International Publishing: Cham, Switzerland, 2017.
28. Goodman, D.; Hofmeister, J.P.; Szidarovszky, F. *Prognostics and Health Management: A Practical Approach to Improving System Reliability Using Condition-Based Data*; John Wiley & Sons: Hoboken, NJ, USA, 2019; ISBN 978-1-119-35670-7.
29. Raouf, I.; Khan, A.; Khalid, S.; Sohail, M.; Azad, M.M.; Kim, H.S. Sensor-Based Prognostic Health Management of Advanced Driver Assistance System for Autonomous Vehicles: A Recent Survey. *Mathematics* **2022**, *10*, 3233. [CrossRef]
30. Baptista, M.; de Medeiros, I.P.; Malere, J.P.; Nascimento, C., Jr.; Prendinger, H.; Henriques, E.M. Comparative case study of life usage and data-driven prognostics techniques using aircraft fault messages. *Comput. Ind.* **2017**, *86*, 1–14. [CrossRef]
31. Kwon, D.; Hodkiewicz, M.R.; Fan, J.; Shibutani, T.; Pecht, M.G. IoT-based prognostics and systems health management for industrial applications. *IEEE Access* **2016**, *4*, 3659–3670. [CrossRef]
32. Liu, Z.; Jia, Z.; Vong, C.-M.; Han, J.; Yan, C.; Pecht, M. A patent analysis of prognostics and health management (PHM) innovations for electrical systems. *IEEE Access* **2018**, *6*, 18088–18107. [CrossRef]
33. Theissler, A.; Pérez-Velázquez, J.; Kettelgerdes, M.; Elger, G. Predictive maintenance enabled by machine learning: Use cases and challenges in the automotive industry. *Reliab. Eng. Syst. Saf.* **2021**, *215*, 107864. [CrossRef]
34. Wong, S.Y.; Chuah, J.H.; Yap, H.J. Technical data-driven tool condition monitoring challenges for CNC milling: A review. *Int. J. Adv. Manuf. Technol.* **2020**, *107*, 4837–4857. [CrossRef]
35. Liu, C.; Cichon, A.; Królczyk, G.; Li, Z. Technology development and commercial applications of industrial fault diagnosis system: A review. *Int. J.Adv. Manuf. Technol.* **2022**, *118*, 3497–3529. [CrossRef]
36. Ranasinghe, K.; Sabatini, R.; Gardi, A.; Bijjahalli, S.; Kapoor, R.; Fahey, T.; Thangavel, K. Advances in Integrated System Health Management for mission-essential and safety-critical aerospace applications. *Prog. Aerosp. Sci.* **2022**, *128*, 100758. [CrossRef]
37. Djurdjanovic, D.; Lee, J.; Ni, J. Watchdog Agent—An infotronics-based prognostics approach for product performance degradation assessment and prediction. *Adv. Eng. Inform.* **2003**, *17*, 109–125. [CrossRef]
38. Nuñez, D.L.; Borsato, M. OntoProg: An ontology-based model for implementing Prognostics Health Management in mechanical machines. *Adv. Eng. Inform.* **2018**, *38*, 746–759. [CrossRef]
39. Gao, R.; Wang, L.; Teti, R.; Dornfeld, D.; Kumara, S.; Mori, M.; Helu, M. Cloud-enabled prognosis for manufacturing. *CIRP Ann.* **2015**, *64*, 749–772. [CrossRef]
40. Pech, M.; Vrchota, J.; Bednář, J. Predictive maintenance and intelligent sensors in smart factory. *Sensors* **2021**, *21*, 1470. [CrossRef] [PubMed]
41. Khalid, S.; Hwang, H.; Kim, H.S. Real-world data-driven machine-learning-based optimal sensor selection approach for equipment fault detection in a thermal power plant. *Mathematics* **2021**, *9*, 2814. [CrossRef]
42. Khalid, S.; Song, J.; Raouf, I.; Kim, H.S. Advances in Fault Detection and Diagnosis for Thermal Power Plants: A Review of Intelligent Techniques. *Mathematics* **2023**, *11*, 1767. [CrossRef]
43. Moore, R. *Making Common Sense Common Practice*; Butterworth–Heinemann: Oxford, UK, 2004.
44. Lin, Y.; Li, X.; Hu, Y. Deep diagnostics and prognostics: An integrated hierarchical learning framework in PHM applications. *Appl. Soft Comput.* **2018**, *72*, 555–564. [CrossRef]
45. Booyse, W.; Wilke, D.N.; Heyns, S. Deep digital twins for detection, diagnostics and prognostics. *Mech. Syst. Signal Process.* **2020**, *140*, 106612. [CrossRef]
46. Qiao, G.; Weiss, B.A. Advancing measurement science to assess monitoring, diagnostics, and prognostics for manufacturing robotics. *Int. J. Progn. Health Manag.* **2016**, *7*, 013. [CrossRef]
47. Xu, J.; Wang, Y.; Xu, L. PHM-oriented integrated fusion prognostics for aircraft engines based on sensor data. *IEEE Sens. J.* **2013**, *14*, 1124–1132. [CrossRef]
48. Raouf, I.; Lee, H.; Kim, H.S. Mechanical fault detection based on machine learning for robotic RV reducer using electrical current signature analysis: A data-driven approach. *J. Comput. Des. Eng.* **2022**, *9*, 417–433. [CrossRef]
49. Siddiqa, A.; Karim, A.; Gani, A. Big data storage technologies: A survey. *Front. Inf. Technol. Electron. Eng.* **2017**, *18*, 1040–1070. [CrossRef]
50. Yang, C.; Huang, Q.; Li, Z.; Liu, K.; Hu, F. Big Data and cloud computing: Innovation opportunities and challenges. *Int. J. Digit. Earth* **2017**, *10*, 13–53. [CrossRef]
51. Nath, A.G.; Udmale, S.S.; Singh, S.K. Role of artificial intelligence in rotor fault diagnosis: A comprehensive review. *Artif Intell Rev* **2021**, *54*, 2609–2668. [CrossRef]
52. Arinez, J.F.; Chang, Q.; Gao, R.X.; Xu, C.; Zhang, J. Artificial Intelligence in Advanced Manufacturing: Current Status and Future Outlook. *J. Manuf. Sci. Eng.* **2020**, *142*, 110804. [CrossRef]
53. Dalzochio, J.; Kunst, R.; Pignaton, E.; Binotto, A.; Sanyal, S.; Favilla, J.; Barbosa, J. Machine learning and reasoning for predictive maintenance in Industry 4.0: Current status and challenges. *Comput. Ind.* **2020**, *123*, 103298. [CrossRef]

54. Alam, M.; Samad, M.D.; Vidyaratne, L.; Glandon, A.; Iftekharuddin, K.M. Survey on deep neural networks in speech and vision systems. *Neurocomputing* **2020**, *417*, 302–321. [CrossRef]
55. Lundervold, A.S.; Lundervold, A. An overview of deep learning in medical imaging focusing on MRI. *Z. Für Med. Phys.* **2019**, *29*, 102–127. [CrossRef]
56. Rohan, A.; Raouf, I.; Kim, H.S. Rotate vector (Rv) reducer fault detection and diagnosis system: Towards component level prognostics and health management (phm). *Sensors* **2020**, *20*, 6845. [CrossRef]
57. Shamshirband, S.; Fathi, M.; Dehzangi, A.; Chronopoulos, A.T.; Alinejad-Rokny, H. A review on deep learning approaches in healthcare systems: Taxonomies, challenges, and open issues. *J. Biomed. Inform.* **2021**, *113*, 103627. [CrossRef]
58. Tobore, I.; Li, J.; Yuhang, L.; Al-Handarish, Y.; Kandwal, A.; Nie, Z.; Wang, L. Deep learning intervention for health care challenges: Some biomedical domain considerations. *JMIR Mhealth Uhealth* **2019**, *7*, e11966. [CrossRef]
59. Yamashita, R.; Nishio, M.; Do, R.K.G.; Togashi, K. Convolutional neural networks: An overview and application in radiology. *Insights Imaging* **2018**, *9*, 611–629. [CrossRef]
60. Kanjo, E.; Younis, E.M.; Ang, C.S. Deep learning analysis of mobile physiological, environmental and location sensor data for emotion detection. *Inf. Fusion* **2019**, *49*, 46–56. [CrossRef]
61. Li, X.; Zhang, W.; Ding, Q. Deep learning-based remaining useful life estimation of bearings using multi-scale feature extraction. *Reliab. Eng. Syst. Saf.* **2019**, *182*, 208–218. [CrossRef]
62. Wang, J.; Ma, Y.; Zhang, L.; Gao, R.X.; Wu, D. Deep learning for smart manufacturing: Methods and applications. *J. Manuf. Syst.* **2018**, *48*, 144–156. [CrossRef]
63. Duan, L.; Xie, M.; Wang, J.; Bai, T. Deep learning enabled intelligent fault diagnosis: Overview and applications. *J. Intell. Fuzzy Syst.* **2018**, *35*, 5771–5784. [CrossRef]
64. Li, G.; Deng, C.; Wu, J.; Chen, Z.; Xu, X. Rolling Bearing Fault Diagnosis Based on Wavelet Packet Transform and Convolutional Neural Network. *Appl. Sci.* **2020**, *10*, 770. [CrossRef]
65. Lee, H.; Raouf, I.; Song, J.; Kim, H.S.; Lee, S. Prognostics and Health Management of the Robotic Servo-Motor under Variable Operating Conditions. *Mathematics* **2023**, *11*, 398. [CrossRef]
66. Raouf, I.; Kumar, P.; Lee, H.; Kim, H.S. Transfer Learning-Based Intelligent Fault Detection Approach for the Industrial Robotic System. *Mathematics* **2023**, *11*, 945. [CrossRef]
67. Zhou, X.; Zhou, H.; He, Y.; Huang, S.; Zhu, Z.; Chen, J. Harmonic reducer in-situ fault diagnosis for industrial robots based on deep learning. *Sci. China Technol. Sci.* **2022**, *65*, 2116–2126. [CrossRef]
68. Adam, H.E.A.; Kimotho, J.K.; Njiri, J.G. Multiple faults diagnosis for an industrial robot fuse quality test bench using deep-learning. *Results Eng.* **2023**, *17*, 101007. [CrossRef]
69. Yin, T.; Lu, N.; Guo, G.; Lei, Y.; Wang, S.; Guan, X. Knowledge and data dual-driven transfer network for industrial robot fault diagnosis. *Mech. Syst. Signal Process.* **2023**, *182*, 109597. [CrossRef]
70. Nagarajan, R. *Introduction to Industrial Robotics*; Pearson Education India: Delhi, India, 2016.
71. Parhami, B. Defect, fault, error, . . . , or failure? *IEEE Trans. Reliab.* **1997**, *46*, 450–451. [CrossRef]
72. Lee, J.; Wu, F.; Zhao, W.; Ghaffari, M.; Liao, L.; Siegel, D. Prognostics and health management design for rotary machinery systems—Reviews, methodology and applications. *Mech. Syst. Signal Process.* **2014**, *42*, 314–334. [CrossRef]
73. Heng, A.; Zhang, S.; Tan, A.C.; Mathew, J. Rotating machinery prognostics: State of the art, challenges and opportunities. *Mech. Syst. Signal Process.* **2009**, *23*, 724–739. [CrossRef]
74. Kumar, P.; Hati, A.S. Review on machine learning algorithm based fault detection in induction motors. *Arch. Comput. Methods Eng.* **2021**, *28*, 1929–1940. [CrossRef]
75. Tavner, P.J. Review of condition monitoring of rotating electrical machines. *IET Electr. Power Appl.* **2008**, *2*, 215–247. [CrossRef]
76. Zhang, L.; Lin, J.; Liu, B.; Zhang, Z.; Yan, X.; Wei, M. A Review on Deep Learning Applications in Prognostics and Health Management. *IEEE Access* **2019**, *7*, 162415–162438. [CrossRef]
77. Gálvez, A.; Diez-Olivan, A.; Seneviratne, D.; Galar, D. Fault Detection and RUL Estimation for Railway HVAC Systems Using a Hybrid Model-Based Approach. *Sustainability* **2021**, *13*, 6828. [CrossRef]
78. Yan, B.; Ma, X.; Huang, G.; Zhao, Y. Two-stage physics-based Wiener process models for online RUL prediction in field vibration data. *Mech. Syst. Signal Process.* **2021**, *152*, 107378. [CrossRef]
79. Liu, H.; Zhou, J.; Xu, Y.; Zheng, Y.; Peng, X.; Jiang, W. Unsupervised fault diagnosis of rolling bearings using a deep neural network based on generative adversarial networks. *Neurocomputing* **2018**, *315*, 412–424. [CrossRef]
80. Cheng, C.; Zhou, B.; Ma, G.; Wu, D.; Yuan, Y. Wasserstein distance based deep adversarial transfer learning for intelligent fault diagnosis with unlabeled or insufficient labeled data. *Neurocomputing* **2020**, *409*, 35–45. [CrossRef]
81. Lu, W.; Liang, B.; Cheng, Y.; Meng, D.; Yang, J.; Zhang, T. Deep Model Based Domain Adaptation for Fault Diagnosis. *IEEE Trans. Ind. Electron.* **2017**, *64*, 2296–2305. [CrossRef]
82. Calabrese, F.; Regattieri, A.; Bortolini, M.; Gamberi, M.; Pilati, F. Predictive Maintenance: A Novel Framework for a Data-Driven, Semi-Supervised, and Partially Online Prognostic Health Management Application in Industries. *Appl. Sci.* **2021**, *11*, 3380. [CrossRef]
83. Rezamand, M.; Kordestani, M.; Carriveau, R.; Ting, D.S.-K.; Orchard, M.E.; Saif, M. Critical Wind Turbine Components Prognostics: A Comprehensive Review. *IEEE Trans. Instrum. Meas.* **2020**, *69*, 9306–9328. [CrossRef]

84. Jin, X.; Sun, Y.; Que, Z.; Wang, Y.; Chow, T.W.S. Anomaly Detection and Fault Prognosis for Bearings. *IEEE Trans. Instrum. Meas.* **2016**, *65*, 2046–2054. [CrossRef]
85. Sikorska, J.Z.; Hodkiewicz, M.; Ma, L. Prognostic modelling options for remaining useful life estimation by industry. *Mech. Syst. Signal Process.* **2011**, *25*, 1803–1836. [CrossRef]
86. Byington, C.; Watson, M.; Sheldon, J.; Swerdon, G. Shaft coupling model-based prognostics enhanced by vibration diagnostics. *Insight-Non-Destr. Test. Cond. Monit.* **2009**, *51*, 420–425. [CrossRef]
87. Fisher, C.; Baines, N. Multi-sensor condition monitoring systems for gas turbines. *J. Cond. Monit.* **1988**, *1*, 57–68.
88. Kemerait, R. New cepstral approach for prognostic maintenance of cyclic machinery. In Proceedings of the IEEE SOUTHEAST-CON, Tampa, FL, USA, 5–8 April 1987; Volume 1987, pp. 256–262.
89. Muir, D.; Taylor, B. Oil debris monitoring for aeroderivative gas turbine. *ASME Power Div. (Publ.) PWR* **1997**, *32*, 547–553.
90. Crow, E.C.; Reichard, K.; Rogan, C.; Callen, J.; Seifert, E. Integrated multi-sensor package (IMSP) for unmanned vehicle operations. In Proceedings of the Unmanned/Unattended Sensors and Sensor Networks IV, International Society for Optics and Photonics. Florence, Italy, 18–20 September 2007; Volume 6736, p. 673604.
91. Liao, L.; Köttig, F. A hybrid framework combining data-driven and model-based methods for system remaining useful life prediction. *Appl. Soft Comput.* **2016**, *44*, 191–199. [CrossRef]
92. Lee, M.-S.; Shifat, T.A.; Hur, J.W. Kalman Filter Assisted Deep Feature Learning for RUL Prediction of Hydraulic Gear Pump. *IEEE Sens. J.* **2022**, *22*, 11088–11097. [CrossRef]
93. Liu, Z.-X.; Wang, Z.-Y.; Wang, Y.; Ji, Z.-C. Optimal Zonotopic Kalman Filter-based State Estimation and Fault-diagnosis Algorithm for Linear Discrete-time System with Time Delay. *Int. J. Control Autom. Syst.* **2022**, *20*, 1757–1771. [CrossRef]
94. Maynard, K.P. Interstitial processing: The application of noise processing to gear fault detection. In Proceedings of the international conference on Condition monitoring, University of Wales, Swansea, UK, 12–16 April 1999; Volume 12, pp. 77–86.
95. Patton, R.J. Fault detection and diagnosis in aerospace systems using analytical redundancy. In Proceedings of the IEE Colloquium on Condition Monitoring and Fault Tolerance, London, UK, 6 November 1990; IET: London, UK, 1990.
96. Tudoroiu, N.; Khorasani, K. Satellite fault diagnosis using a bank of interacting Kalman filters. *IEEE Trans. Aerosp. Electron. Syst.* **2007**, *43*, 1334–1350. [CrossRef]
97. Hansen, R.J.; Hall, D.L.; Kurtz, S.K. A new approach to the challenge of machinery prognostics. In Proceedings of the Turbo Expo: Power for Land, Sea, and Air, The Hague, The Netherlands, 13–16 June 1994; American Society of Mechanical Engineers: New York, NY, USA, 1994; Volume 78873, p. V005T15A001.
98. Denkena, B.; Litwinski, K.; Brouwer, D.; Boujnah, H. Design and analysis of a prototypical sensory Z-slide for machine tools. *Prod. Eng.* **2013**, *7*, 9–14. [CrossRef]
99. Holland, S.W.; Barajas, L.G.; Salman, M.; Zhang, Y. PHM for automotive manufacturing & vehicle applications. In Proceedings of the Prognostics & Health Management Conference, Portland, OR, USA, 10–14 October 2010.
100. Hu, S.J.; Koren, Y. Stream-of-variation theory for automotive body assembly. *CIRP Ann.* **1997**, *46*, 1–6. [CrossRef]
101. Sahir Arıkan, M.; Balkan, T. Process modeling, simulation, and paint thickness measurement for robotic spray painting. *J. Robot. Syst.* **2000**, *17*, 479–494. [CrossRef]
102. Shen, T.; Wan, F.; Cui, W.; Song, B. Application of prognostic and health management technology on aircraft fuel system. In Proceedings of the 2010 Prognostics and System Health Management Conference, Macau, China, 12–14 January 2010; IEEE: New York, NY, USA, 2010; pp. 1–7.
103. Siegel, D.; Zhao, W.; Lapira, E.; AbuAli, M.; Lee, J. A comparative study on vibration-based condition monitoring algorithms for wind turbine drive trains. *Wind Energy* **2014**, *17*, 695–714. [CrossRef]
104. Siegel, D.; Lee, J.; Dempsey, P. Investigation and Evaluation of Condition Indicators, Variable Selection, and Health Indication Methods and Algorithms For Rotorcraft Gear Components. In Proceedings of the MFPT 2014 Conference, Virginia Beach, VA, USA, 10–12 June 2014.
105. Bin, G.; Gao, J.; Li, X.; Dhillon, B. Early fault diagnosis of rotating machinery based on wavelet packets—Empirical mode decomposition feature extraction and neural network. *Mech. Syst. Signal Process.* **2012**, *27*, 696–711. [CrossRef]
106. Li, W.; Zhu, Z.; Jiang, F.; Zhou, G.; Chen, G. Fault diagnosis of rotating machinery with a novel statistical feature extraction and evaluation method. *Mech. Syst. Signal Process.* **2015**, *50*, 414–426. [CrossRef]
107. Zhu, H.; He, Z.; Wei, J.; Wang, J.; Zhou, H. Bearing fault feature extraction and fault diagnosis method based on feature fusion. *Sensors* **2021**, *21*, 2524. [CrossRef]
108. Widodo, A.; Yang, B.-S. Application of nonlinear feature extraction and support vector machines for fault diagnosis of induction motors. *Expert Syst. Appl.* **2007**, *33*, 241–250. [CrossRef]
109. Cheng, G.; Chen, X.; Li, H.; Li, P.; Liu, H. Study on planetary gear fault diagnosis based on entropy feature fusion of ensemble empirical mode decomposition. *Measurement* **2016**, *91*, 140–154. [CrossRef]
110. Zimroz, R.; Bartkowiak, A. Two simple multivariate procedures for monitoring planetary gearboxes in non-stationary operating conditions. *Mech. Syst. Signal Process.* **2013**, *38*, 237–247. [CrossRef]
111. Kumar, P.; Hati, A.S. Support Vector Classifier-Based Broken Rotor Bar Detection in Squirrel Cage Induction Motor. In *Machines, Mechanism and Robotics*; Springer: Berlin/Heidelberg, Germany, 2022; pp. 429–438.
112. Moghaddass, R.; Sheng, S. An anomaly detection framework for dynamic systems using a Bayesian hierarchical framework. *Appl. Energy* **2019**, *240*, 561–582. [CrossRef]

113. Cabrera, D.; Sancho, F.; Sánchez, R.-V.; Zurita, G.; Cerrada, M.; Li, C.; Vásquez, R.E. Fault diagnosis of spur gearbox based on random forest and wavelet packet decomposition. *Front. Mech. Eng.* **2015**, *10*, 277–286. [CrossRef]
114. Tian, J.; Morillo, C.; Azarian, M.H.; Pecht, M. Motor bearing fault detection using spectral kurtosis-based feature extraction coupled with K-nearest neighbor distance analysis. *IEEE Trans. Ind. Electron.* **2015**, *63*, 1793–1803. [CrossRef]
115. Chen, J.; Ma, C.; Song, D.; Xu, B. Failure prognosis of multiple uncertainty system based on Kalman filter and its application to aircraft fuel system. *Adv. Mech. Eng.* **2016**, *8*, 1–13. [CrossRef]
116. Singleton, R.K.; Strangas, E.G.; Aviyente, S. Extended Kalman filtering for remaining-useful-life estimation of bearings. *IEEE Trans. Ind. Electron.* **2014**, *62*, 1781–1790. [CrossRef]
117. Soualhi, A.; Clerc, G.; Razik, H.; Guillet, F. Hidden Markov models for the prediction of impending faults. *IEEE Trans. Ind. Electron.* **2016**, *63*, 3271–3281. [CrossRef]
118. Niu, G. *Data-Driven Technology for Engineering Systems Health Management*; Springer: Berlin/Heidelberg, Germany, 2017.
119. Aissani, N.; Beldjilali, B.; Trentesaux, D. Dynamic scheduling of maintenance tasks in the petroleum industry: A reinforcement approach. *Eng. Appl. Artif. Intell.* **2009**, *22*, 1089–1103. [CrossRef]
120. Chan, G.; Asgarpoor, S. Optimum maintenance policy with Markov processes. *Electr. Power Syst. Res.* **2006**, *76*, 452–456. [CrossRef]
121. Wu, S.; Gebraeel, N.; Lawley, M.A.; Yih, Y. A neural network integrated decision support system for condition-based optimal predictive maintenance policy. *IEEE Trans. Syst. Man Cybern.-Part A Syst. Hum.* **2007**, *37*, 226–236. [CrossRef]
122. Ali, J.B.; Fnaiech, N.; Saidi, L.; Chebel-Morello, B.; Fnaiech, F. Application of empirical mode decomposition and artificial neural network for automatic bearing fault diagnosis based on vibration signals. *Appl. Acoust.* **2015**, *89*, 16–27.
123. Zhu, J.; Chen, N.; Peng, W. Estimation of bearing remaining useful life based on multiscale convolutional neural network. *IEEE Trans. Ind. Electron.* **2018**, *66*, 3208–3216. [CrossRef]
124. Deutsch, J.; He, D. Using deep learning-based approach to predict remaining useful life of rotating components. *IEEE Trans. Syst. Man Cybern. Syst.* **2017**, *48*, 11–20. [CrossRef]
125. Shao, H.; Jiang, H.; Lin, Y.; Li, X. A novel method for intelligent fault diagnosis of rolling bearings using ensemble deep auto-encoders. *Mech. Syst. Signal Process.* **2018**, *102*, 278–297. [CrossRef]
126. Kumar, P.; Hati, A.S. Deep convolutional neural network based on adaptive gradient optimizer for fault detection in SCIM. *ISA Trans.* **2021**, *111*, 350–359. [CrossRef]
127. Saxena, A.; Celaya, J.; Saha, B.; Saha, S.; Goebel, K. Metrics for offline evaluation of prognostic performance. *Int. J. Progn. Health Manag.* **2010**, *1*, 4–23. [CrossRef]
128. Lou, X.; Loparo, K.A. Bearing fault diagnosis based on wavelet transform and fuzzy inference. *Mech. Syst. Signal Process.* **2004**, *18*, 1077–1095. [CrossRef]
129. Batista, L.; Badri, B.; Sabourin, R.; Thomas, M. A classifier fusion system for bearing fault diagnosis. *Expert Syst. Appl.* **2013**, *40*, 6788–6797. [CrossRef]
130. Chen, N.; Tsui, K.L. Condition monitoring and remaining useful life prediction using degradation signals: Revisited. *IIE Trans.* **2013**, *45*, 939–952. [CrossRef]
131. Nectoux, P.; Gouriveau, R.; Medjaher, K.; Ramasso, E.; Chebel-Morello, B.; Zerhouni, N.; Varnier, C. PRONOSTIA: An experimental platform for bearings accelerated degradation tests. In Proceedings of the IEEE International Conference on Prognostics and Health Management, PHM'12, Paris, France, 31 May–2 June 2012; IEEE Catalog Number: CPF12PHM-CDR. 2012; pp. 1–8.
132. Erfani, S.M.; Rajasegarar, S.; Karunasekera, S.; Leckie, C. High-dimensional and large-scale anomaly detection using a linear one-class SVM with deep learning. *Pattern Recognit.* **2016**, *58*, 121–134. [CrossRef]
133. Chen, S.; Yu, J.; Wang, S. One-dimensional convolutional auto-encoder-based feature learning for fault diagnosis of multivariate processes. *J. Process Control* **2020**, *87*, 54–67. [CrossRef]
134. Givnan, S.; Chalmers, C.; Fergus, P.; Ortega-Martorell, S.; Whalley, T. Anomaly Detection Using Autoencoder Reconstruction upon Industrial Motors. *Sensors* **2022**, *22*, 3166. [CrossRef]
135. Yang, Z.; Xu, B.; Luo, W.; Chen, F. Autoencoder-based representation learning and its application in intelligent fault diagnosis: A review. *Measurement* **2022**, *189*, 110460. [CrossRef]
136. Peyron, M.; Fillion, A.; Gürol, S.; Marchais, V.; Gratton, S.; Boudier, P.; Goret, G. Latent space data assimilation by using deep learning. *Q. J. R. Meteorol. Soc.* **2021**, *147*, 3759–3777. [CrossRef]
137. d'Acremont, A.; Fablet, R.; Baussard, A.; Quin, G. CNN-Based Target Recognition and Identification for Infrared Imaging in Defense Systems. *Sensors* **2019**, *19*, 2040. [CrossRef]
138. Shi, Y.; Yang, J.; Qi, Z. Unsupervised anomaly segmentation via deep feature reconstruction. *Neurocomputing* **2021**, *424*, 9–22. [CrossRef]
139. Karabacak, Y.E.; Gürsel Özmen, N. Common spatial pattern-based feature extraction and worm gear fault detection through vibration and acoustic measurements. *Measurement* **2022**, *187*, 110366. [CrossRef]
140. Wen, Y.; Fashiar Rahman, M.; Xu, H.; Tseng, T.-L.B. Recent advances and trends of predictive maintenance from data-driven machine prognostics perspective. *Measurement* **2022**, *187*, 110276. [CrossRef]
141. Zhang, J.; Zeng, Y.; Starly, B. Recurrent neural networks with long term temporal dependencies in machine tool wear diagnosis and prognosis. *SN Appl. Sci.* **2021**, *3*, 442. [CrossRef]

142. Chen, C.; Zhu, Z.H.; Shi, J.; Lu, N.; Jiang, B. Dynamic Predictive Maintenance Scheduling Using Deep Learning Ensemble for System Health Prognostics. *IEEE Sens. J.* **2021**, *21*, 26878–26891. [CrossRef]
143. Goodfellow, I.; Bengio, Y.; Courville, A. *Deep Learning*; MIT Press: Cambridge, MA, USA, 2016.
144. Dargan, S.; Kumar, M.; Ayyagari, M.R.; Kumar, G. A Survey of Deep Learning and Its Applications: A New Paradigm to Machine Learning. *Arch. Comput. Methods Eng* **2020**, *27*, 1071–1092. [CrossRef]
145. Arel, I.; Rose, D.C.; Karnowski, T.P. Deep Machine Learning—A New Frontier in Artificial Intelligence Research [Research Frontier]. *IEEE Comput. Intell. Mag.* **2010**, *5*, 13–18. [CrossRef]
146. Aggarwal, C.C. *Neural Networks and Deep Learning: A Textbook*; Springer: Berlin/Heidelberg, Germany, 2018; ISBN 978-3-319-94463-0.
147. Alzubaidi, L.; Zhang, J.; Humaidi, A.J.; Al-Dujaili, A.; Duan, Y.; Al-Shamma, O.; Santamaría, J.; Fadhel, M.A.; Al-Amidie, M.; Farhan, L. Review of deep learning: Concepts, CNN architectures, challenges, applications, future directions. *J. Big Data* **2021**, *8*, 53. [CrossRef] [PubMed]
148. Shrestha, A.; Mahmood, A. Review of Deep Learning Algorithms and Architectures. *IEEE Access* **2019**, *7*, 53040–53065. [CrossRef]
149. Hinton, G.E. Deep belief networks. *Scholarpedia* **2009**, *4*, 5947. [CrossRef]
150. Ng, A. Sparse autoencoder. *CS294A Lect. Notes* **2011**, *72*, 1–19.
151. Vincent, P.; Larochelle, H.; Lajoie, I.; Bengio, Y.; Manzagol, P.-A.; Bottou, L. Stacked denoising autoencoders: Learning useful representations in a deep network with a local denoising criterion. *J. Mach. Learn. Res.* **2010**, *11*, 3371–3408.
152. An, J.; Cho, S. Variational autoencoder based anomaly detection using reconstruction probability. *Spec. Lect. IE* **2015**, *2*, 1–18.
153. Le Cun, Y.; Boser, B.; Denker, J.S.; Henderson, D.; Howard, R.E.; Hubbard, W.; Jackel, L.D. Handwritten digit recognition with a back-propagation network. In Proceedings of the 2nd International Conference on Neural Information Processing Systems, Denver, CO, USA, 27–30 November 1989; pp. 396–404.
154. Dash, P.B.; Naik, B.; Nayak, J.; Vimal, S. Deep belief network-based probabilistic generative model for detection of robotic manipulator failure execution. *Soft Comput.* **2021**, *27*, 363–375. [CrossRef]
155. Chen, Z.; Li, W. Multisensor feature fusion for bearing fault diagnosis using sparse autoencoder and deep belief network. *IEEE Trans. Instrum. Meas.* **2017**, *66*, 1693–1702. [CrossRef]
156. Ren, H.; Chai, Y.; Qu, J.; Ye, X.; Tang, Q. A novel adaptive fault detection methodology for complex system using deep belief networks and multiple models: A case study on cryogenic propellant loading system. *Neurocomputing* **2018**, *275*, 2111–2125. [CrossRef]
157. Xing, S.; Lei, Y.; Wang, S.; Jia, F. Distribution-invariant deep belief network for intelligent fault diagnosis of machines under new working conditions. *IEEE Trans. Ind. Electron.* **2020**, *68*, 2617–2625. [CrossRef]
158. Jiao, J.; Zheng, X. Fault Diagnosis Method for Industrial Robots Based on DBN Joint Information Fusion Technology. *Comput. Intell. Neurosci.* **2022**, *2022*, 4340817. [CrossRef] [PubMed]
159. Ji, C.; Wang, K.; Yuan, D. Health assessment method of industrial robot reducer based on deep belief network. In Proceedings of the 2021 IEEE 11th Annual International Conference on CYBER Technology in Automation, Control, and Intelligent Systems (CYBER), Jiaxing, China, 27–31 July 2021; IEEE: New York, NY, USA, 2021; pp. 442–446.
160. Shao, S.-Y.; Sun, W.-J.; Yan, R.-Q.; Wang, P.; Gao, R.X. A deep learning approach for fault diagnosis of induction motors in manufacturing. *Chin. J. Mech. Eng.* **2017**, *30*, 1347–1356. [CrossRef]
161. Hu, G.; Li, H.; Xia, Y.; Luo, L. A deep Boltzmann machine and multi-grained scanning forest ensemble collaborative method and its application to industrial fault diagnosis. *Comput. Ind.* **2018**, *100*, 287–296. [CrossRef]
162. Deng, S.; Cheng, Z.; Li, C.; Yao, X.; Chen, Z.; Sanchez, R.-V. Rolling bearing fault diagnosis based on Deep Boltzmann machines. In Proceedings of the 2016 Prognostics and System Health Management Conference (PHM-Chengdu), Chengdu, China, 19–21 October 2016; IEEE: New York, NY, USA, 2016; pp. 1–6.
163. Li, C.; Sanchez, R.-V.; Zurita, G.; Cerrada, M.; Cabrera, D.; Vásquez, R.E. Multimodal deep support vector classification with homologous features and its application to gearbox fault diagnosis. *Neurocomputing* **2015**, *168*, 119–127. [CrossRef]
164. Wang, J.; Wang, K.; Wang, Y.; Huang, Z.; Xue, R. Deep Boltzmann machine based condition prediction for smart manufacturing. *J. Ambient Intell. Humaniz. Comput.* **2019**, *10*, 851–861. [CrossRef]
165. Mitrevski, A.; Plöger, P.G. Data-Driven Robot Fault Detection and Diagnosis Using Generative Models: A Modified SFDD Algorithm. In Proceedings of the 30th International Workshop on Principles of Diagnosis DX'19, Kla-genfurt, Austria, 11–13 November 2019.
166. Chen, T.; Liu, X.; Xia, B.; Wang, W.; Lai, Y. Unsupervised anomaly detection of industrial robots using sliding-window convolutional variational autoencoder. *IEEE Access* **2020**, *8*, 47072–47081. [CrossRef]
167. Hong, Y.; Sun, Z.; Zou, X.; Long, J. Multi-joint Industrial Robot Fault Identification using Deep Sparse Auto-Encoder Network with Attitude Data. In Proceedings of the 2020 Prognostics and Health Management Conference (PHM-Besançon), Besancon, France, 4–7 May 2020; IEEE: New York, NY, USA, 2020; pp. 176–179.
168. 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* **2021**, *176*, 109186. [CrossRef]
169. Yun, H.; Kim, H.; Jeong, Y.H.; Jun, M.B. Autoencoder-based anomaly detection of industrial robot arm using stethoscope based internal sound sensor. *J. Intell. Manuf.* **2021**, *34*, 1427–1444. [CrossRef]

170. Sun, W.; Shao, S.; Zhao, R.; Yan, R.; Zhang, X.; Chen, X. A sparse auto-encoder-based deep neural network approach for induction motor faults classification. *Measurement* **2016**, *89*, 171–178. [CrossRef]
171. Li, C.; Zhang, W.; Peng, G.; Liu, S. Bearing fault diagnosis using fully-connected winner-take-all autoencoder. *IEEE Access* **2017**, *6*, 6103–6115. [CrossRef]
172. Sohaib, M.; Kim, J.-M. Reliable fault diagnosis of rotary machine bearings using a stacked sparse autoencoder-based deep neural network. *Shock Vib.* **2018**, *2018*, 2919637. [CrossRef]
173. Chen, L.; Cao, J.; Wu, K.; Zhang, Z. Application of Generalized Frequency Response Functions and Improved Convolutional Neural Network to Fault Diagnosis of Heavy-duty Industrial Robot. *Robot. Comput.-Integr. Manuf.* **2022**, *73*, 102228. [CrossRef]
174. Oh, Y.; Kim, Y.; Na, K.; Youn, B.D. A deep transferable motion-adaptive fault detection method for industrial robots using a residual–convolutional neural network. *ISA Trans.* **2021**, *128*, 521–534. [CrossRef]
175. Yang, S.; Luo, X.; Li, C. Fault Diagnosis of Rotation Vector Reducer for Industrial Robot Based on a Convolutional Neural Network. *Stroj. Vestn. /J. Mech. Eng.* **2021**, *67*, 489–500. [CrossRef]
176. Ma, Z.; Xiao, H.; Pan, Y.; Jiang, W.; Xiong, M.; He, Z. Multi-axis Industrial Robot Fault Diagnosis Model Based on Improved One-Dimensional Convolutional Neural Network. In Proceedings of the International Conference on Data Mining and Big Data, Guangzhou, China, 20–22 October 2021; pp. 397–410.
177. Li, P.; Xiao, H.; Jiang, W.; Ning, D. Compound Fault Diagnosis of Industrial Robot Based on Improved Multi-label One-Dimensional Convolutional Neural Network. In Proceedings of the International Conference on Data Mining and Big Data, Guangzhou, China, 20–22 October 2021; Springer: Berlin/Heidelberg, Germany, 2021; pp. 205–216.
178. Liu, Y.; Chen, C.; Wang, T.; Cheng, L. An attention enhanced dilated CNN approach for cross-axis industrial robotics fault diagnosis. *Auton. Intell. Syst.* **2022**, *2*, 11. [CrossRef]
179. Lu, K.; Chen, C.; Wang, T.; Cheng, L.; Qin, J. Fault diagnosis of industrial robot based on dual-module attention convolutional neural network. *Auton. Intell. Syst.* **2022**, *2*, 12. [CrossRef]
180. Janssens, O.; Slavkovikj, V.; Vervisch, B.; Stockman, K.; Loccufier, M.; Verstockt, S.; Van de Walle, R.; Van Hoecke, S. Convolutional neural network based fault detection for rotating machinery. *J. Sound Vib.* **2016**, *377*, 331–345. [CrossRef]
181. Plakias, S.; Boutalis, Y.S. Fault detection and identification of rolling element bearings with Attentive Dense CNN. *Neurocomputing* **2020**, *405*, 208–217. [CrossRef]
182. Cheng, Y.; Lin, M.; Wu, J.; Zhu, H.; Shao, X. Intelligent fault diagnosis of rotating machinery based on continuous wavelet transform-local binary convolutional neural network. *Knowl.-Based Syst.* **2021**, *216*, 106796. [CrossRef]
183. Guo, S.; Yang, T.; Gao, W.; Zhang, C. A novel fault diagnosis method for rotating machinery based on a convolutional neural network. *Sensors* **2018**, *18*, 1429. [CrossRef] [PubMed]
184. Liang, P.; Deng, C.; Wu, J.; Yang, Z. Intelligent fault diagnosis of rotating machinery via wavelet transform, generative adversarial nets and convolutional neural network. *Measurement* **2020**, *159*, 107768. [CrossRef]
185. Li, Y.; Zou, L.; Jiang, L.; Zhou, X. Fault diagnosis of rotating machinery based on combination of deep belief network and one-dimensional convolutional neural network. *IEEE Access* **2019**, *7*, 165710–165723. [CrossRef]
186. An, Z.; Li, S.; Wang, J.; Jiang, X. A novel bearing intelligent fault diagnosis framework under time-varying working conditions using recurrent neural network. *ISA Trans.* **2020**, *100*, 155–170. [CrossRef]
187. Van der Maaten, L.; Hinton, G. Visualizing data using t-SNE. *J. Mach. Learn. Res.* **2008**, *9*, 2579–2605.
188. Zhang, Y.; Zhou, T.; Huang, X.; Cao, L.; Zhou, Q. Fault diagnosis of rotating machinery based on recurrent neural networks. *Measurement* **2021**, *171*, 108774. [CrossRef]
189. Liu, H.; Zhou, J.; Zheng, Y.; Jiang, W.; Zhang, Y. Fault diagnosis of rolling bearings with recurrent neural network-based autoencoders. *ISA Trans.* **2018**, *77*, 167–178. [CrossRef]
190. Jiang, H.; Li, X.; Shao, H.; Zhao, K. Intelligent fault diagnosis of rolling bearings using an improved deep recurrent neural network. *Meas. Sci. Technol.* **2018**, *29*, 065107. [CrossRef]
191. Qiao, M.; Yan, S.; Tang, X.; Xu, C. Deep convolutional and LSTM recurrent neural networks for rolling bearing fault diagnosis under strong noises and variable loads. *IEEE Access* **2020**, *8*, 66257–66269. [CrossRef]
192. Oh, S.; Han, S.; Jeong, J. Multi-scale convolutional recurrent neural network for bearing fault detection in noisy manufacturing environments. *Appl. Sci.* **2021**, *11*, 3963. [CrossRef]
193. Li, S.; Zhao, Y.; Ding, M. Mobile robot motor bearing fault detection and classification on discrete wavelet transform and lstm network. *J. Mech. Med. Biol.* **2018**, *18*, 1840034. [CrossRef]
194. Zhi, Z.; Liu, L.; Liu, D.; Hu, C. Fault Detection of the Harmonic Reducer Based on CNN-LSTM With a Novel Denoising Algorithm. *IEEE Sens. J.* **2022**, *22*, 2572–2581. [CrossRef]
195. Wang, T.; Zhang, L.; Wang, X. Fault detection for motor drive control system of industrial robots using CNN-LSTM-based observers. *CES Trans. Electr. Mach. Syst.* **2023**, 1–9. [CrossRef]
196. Sohaib, M.; Kim, C.-H.; Kim, J.-M. A Hybrid Feature Model and Deep-Learning-Based Bearing Fault Diagnosis. *Sensors* **2017**, *17*, 2876. [CrossRef]
197. Xu, Y.; Li, Z.; Wang, S.; Li, W.; Sarkodie-Gyan, T.; Feng, S. A hybrid deep-learning model for fault diagnosis of rolling bearings. *Measurement* **2021**, *169*, 108502. [CrossRef]
198. Hoang, D.T.; Kang, H.J. A Motor Current Signal-Based Bearing Fault Diagnosis Using Deep Learning and Information Fusion. *IEEE Trans. Instrum. Meas.* **2020**, *69*, 3325–3333. [CrossRef]

199. Li, X.; Zhang, W.; Ding, Q. Understanding and improving deep learning-based rolling bearing fault diagnosis with attention mechanism. *Signal Process.* **2019**, *161*, 136–154. [CrossRef]
200. Chen, Z.; Chen, X.; Li, C.; Sanchez, R.-V.; Qin, H. Vibration-based gearbox fault diagnosis using deep neural networks. *J. Vibroengineering* **2017**, *19*, 2475–2496. [CrossRef]
201. Chen, Z.; Mauricio, A.; Li, W.; Gryllias, K. A deep learning method for bearing fault diagnosis based on Cyclic Spectral Coherence and Convolutional Neural Networks. *Mech. Syst. Signal Process.* **2020**, *140*, 106683. [CrossRef]
202. Verstraete, D.; Ferrada, A.; Droguett, E.L.; Meruane, V.; Modarres, M. Deep Learning Enabled Fault Diagnosis Using Time-Frequency Image Analysis of Rolling Element Bearings. *Shock Vib.* **2017**, *2017*, e5067651. [CrossRef]
203. Liu, H.; Li, L.; Ma, J. Rolling Bearing Fault Diagnosis Based on STFT-Deep Learning and Sound Signals. *Shock Vib.* **2016**, *2016*, e6127479. [CrossRef]
204. Shi, J.; Peng, D.; Peng, Z.; Zhang, Z.; Goebel, K.; Wu, D. Planetary gearbox fault diagnosis using bidirectional-convolutional LSTM networks. *Mech. Syst. Signal Process.* **2022**, *162*, 107996. [CrossRef]
205. Ravikumar, K.N.; Yadav, A.; Kumar, H.; Gangadharan, K.V.; Narasimhadhan, A.V. Gearbox fault diagnosis based on Multi-Scale deep residual learning and stacked LSTM model. *Measurement* **2021**, *186*, 110099. [CrossRef]
206. Long, J.; Mou, J.; Zhang, L.; Zhang, S.; Li, C. Attitude data-based deep hybrid learning architecture for intelligent fault diagnosis of multi-joint industrial robots. *J. Manuf. Syst.* **2021**, *61*, 736–745. [CrossRef]
207. Kamat, P.V.; Sugandhi, R.; Kumar, S. Deep learning-based anomaly-onset aware remaining useful life estimation of bearings. *PeerJ Comput. Sci.* **2021**, *7*, e795. [CrossRef]
208. Wang, X.; Wang, T.; Ming, A.; Han, Q.; Chu, F.; Zhang, W.; Li, A. Deep Spatiotemporal Convolutional-Neural-Network-Based Remaining Useful Life Estimation of Bearings. *Chin. J. Mech. Eng.* **2021**, *34*, 62. [CrossRef]
209. Ding, H.; Yang, L.; Cheng, Z.; Yang, Z. A remaining useful life prediction method for bearing based on deep neural networks. *Measurement* **2021**, *172*, 108878. [CrossRef]
210. Li, H.; Wang, W.; Li, Z.; Dong, L.; Li, Q. A novel approach for predicting tool remaining useful life using limited data. *Mech. Syst. Signal Process.* **2020**, *143*, 106832. [CrossRef]
211. Qin, Y.; Xiang, S.; Chai, Y.; Chen, H. Macroscopic–Microscopic Attention in LSTM Networks Based on Fusion Features for Gear Remaining Life Prediction. *IEEE Trans. Ind. Electron.* **2020**, *67*, 10865–10875. [CrossRef]
212. Algburi, R.N.A.; Gao, H. Health Assessment and Fault Detection System for an Industrial Robot Using the Rotary Encoder Signal. *Energies* **2019**, *12*, 2816. [CrossRef]
213. Jaber, A.; Bicker, R. Industrial Robot Fault Detection Based on Wavelet Transform and LabVIEW. In Proceedings of the 2014 First International Conference on Systems Informatics, Modelling and Simulation, Sheffield, UK, 29 April–1 May 2014.
214. Jaber, A.A.; Bicker, R. The optimum selection of wavelet transform parameters for the purpose of fault detection in an industrial robot. In Proceedings of the 2014 IEEE International Conference on Control System, Computing and Engineering (ICCSCE 2014), Penang, Malaysia, 28–30 November 2014; pp. 304–309.
215. Zhang, Y.; An, H.; Ding, X.; Liang, W.; Yuan, M.; Ji, C.; Tan, J. Industrial Robot Rotate Vector Reducer Fault Detection Based on Hidden Markov Models. In Proceedings of the 2019 IEEE International Conference on Robotics and Biomimetics (ROBIO), Dali, China, 6–8 December 2019; pp. 3013–3018.
216. Bonci, A.; Longhi, S.; Nabissi, G.; Verdini, F. Predictive Maintenance System using motor current signal analysis for Industrial Robot. In Proceedings of the 2019 24th IEEE International Conference on Emerging Technologies and Factory Automation (ETFA), Zaragoza, Spain, 10–13 September 2019; pp. 1453–1456.
217. Jaber, A.A.; Bicker, R. Fault diagnosis of industrial robot gears based on discrete wavelet transform and artificial neural network. *Insight-Non-Destr. Test. Cond. Monit.* **2016**, *58*, 179–186. [CrossRef]
218. Chen, Z.; Wu, K.; Wu, J.; Deng, C.; Wang, Y. Residual shrinkage transformer relation network for intelligent fault detection of industrial robot with zero-fault samples. *Knowl.-Based Syst.* **2023**, *268*, 110452. [CrossRef]
219. Jaber, A.A.; Bicker, R. Industrial Robot Backlash Fault Diagnosis Based on Discrete Wavelet Transform and Artificial Neural Network. *Am. J. Mech. Eng.* **2016**, *4*, 21–31. [CrossRef]
220. Yang, Q.; Li, X.; Wang, Y.; Ainapure, A.; Lee, J. Fault Diagnosis of Ball Screw in Industrial Robots Using Non-Stationary Motor Current Signals. *Procedia Manuf.* **2020**, *48*, 1102–1108. [CrossRef]
221. Jaber, A.A.; Bicker, R. Development of a Condition Monitoring Algorithm for Industrial Robots based on Artificial Intelligence and Signal Processing Techniques. *Int. J. Electr. Comput. Eng. (IJECE)* **2018**, *8*, 996–1009. [CrossRef]
222. Nentwich, C.; Reinhart, G. A Method for Health Indicator Evaluation for Condition Monitoring of Industrial Robot Gears. *Robotics* **2021**, *10*, 80. [CrossRef]
223. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep residual learning for image recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016; pp. 770–778.
224. Simonyan, K.; Zisserman, A. Very deep convolutional networks for large-scale image recognition. *arXiv* **2014**, arXiv:1409.1556.
225. Szegedy, C.; Ioffe, S.; Vanhoucke, V.; Alemi, A.A. Inception-v4, inception-resnet and the impact of residual connections on learning. In Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence, San Francisco, CA, USA, 4–9 February 2017.
226. Li, X.; Zhang, W.; Ding, Q.; Sun, J.-Q. Intelligent rotating machinery fault diagnosis based on deep learning using data augmentation. *J. Intell. Manuf.* **2020**, *31*, 433–452. [CrossRef]

227. Kumar, P.; Hati, A.S. Transfer learning-based deep CNN model for multiple faults detection in SCIM. *Neural Comput. Appl.* **2021**, *33*, 15851–15862. [CrossRef]
228. Shao, S.; McAleer, S.; Yan, R.; Baldi, P. Highly accurate machine fault diagnosis using deep transfer learning. *IEEE Trans. Ind. Inform.* **2018**, *15*, 2446–2455. [CrossRef]