

## Article

# Fault Classification for Cooling System of Hydraulic Machinery Using AI

Haseeb Ahmed Khan <sup>1</sup>, Uzair Bhatti <sup>1,\*</sup>, Khurram Kamal <sup>1</sup>, Mohammed Alkahtani <sup>2</sup>, Mustufa Haider Abidi <sup>2</sup>  
and Senthan Mathavan <sup>3</sup>

<sup>1</sup> Department of Engineering Sciences, National University of Sciences and Technology, Islamabad 44000, Pakistan; haseebahmedkhan.92@gmail.com (H.A.K.); khurram.kamal@pnec.nust.edu.pk (K.K.)

<sup>2</sup> Department of Industrial Engineering, College of Engineering, King Saud University, P.O. Box 800, Riyadh 11421, Saudi Arabia; moalkahtani@ksu.edu.sa (M.A.); mabidi@ksu.edu.sa (M.H.A.)

<sup>3</sup> Department of Civil and Structural Engineering, Nottingham Trent University, Burton Street, Nottingham NG1 4BU, UK; b00021159@gmail.com

\* Correspondence: buzair6@gmail.com

**Abstract:** Hydraulic systems are used in all kinds of industries. Mills, manufacturing, robotics, and Ports require the use of Hydraulic Equipment. Many industries prefer to use hydraulic systems due to their numerous advantages over electrical and mechanical systems. Hence, the growth in demand for hydraulic systems has been increasing over time. Due to its vast variety of applications, the faults in hydraulic systems can cause a breakdown. Using Artificial-Intelligence (AI)-based approaches, faults can be classified and predicted to avoid downtime and ensure sustainable operations. This research work proposes a novel approach for the classification of the cooling behavior of a hydraulic test rig. Three fault conditions for the cooling system of the hydraulic test rig were used. The spectrograms were generated using the time series data for three fault conditions. The CNN variant, the Residual Network, was used for the classification of the fault conditions. Various features were extracted from the data including the F-score, precision, accuracy, and recall using a Confusion Matrix. The data contained 43,680 attributes and 2205 instances. After testing, validating, and training, the model accuracy of the ResNet-18 architecture was found to be close to 95%.

**Keywords:** hydraulic systems; breakdown; Artificial Intelligence; sustainable; fault conditions; spectrograms; hydraulic test rig; sensors; machine learning; deep learning



**Citation:** Khan, H.A.; Bhatti, U.; Kamal, K.; Alkahtani, M.; Abidi, M.H.; Mathavan, S. Fault Classification for Cooling System of Hydraulic Machinery Using AI. *Sensors* **2023**, *23*, 7152. <https://doi.org/10.3390/s23167152>

Academic Editor: Davide Astolfi

Received: 22 June 2023

Revised: 3 August 2023

Accepted: 9 August 2023

Published: 13 August 2023



**Copyright:** © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

According to a leading market company for global coverage, technavio, the global Hydraulic Equipment Market will accelerate its growth at a Compound Annual Growth Rate (CAGR) of 4.71% for the period of 2020–2025. It had 3.23% growth for the year 2021. The growth contributed by the Asia-Pacific (APAC) Region is 47%. The incremental growth during the period 2020–2025 is estimated to be around USD 15.50 billion [1]. Hydraulic systems are used in almost every industrial sector, and their application is widely known in major industries including manufacturing, construction, and robotics, to name a few. They are notable in industries for a reason: cost-effectiveness, efficacy, and adaptability. Many industries prefer to use hydraulic systems due to their numerous advantages over mechanical, electrical, and pneumatic systems. Their capability of moving heavier loads with sustained force and torque and providing significantly more power than electrical, mechanical, and pneumatic systems are not surprising to anyone. The advantage of a fluid power system is that it can easily get through a range of heavy loads without having to use gears, levers, and pulleys.

The increasing demand for Hydraulic Equipment requires flawless operation, and finding the source of the problem becomes a challenging task when the hydraulic system

fails. This may include the motor, pump, valves, actuators, and hydraulic fluid. Human error and faulty maintenance practices can be additional sources of failure. Some of the common reasons for hydraulic failure include water and air pollution, temperature problems, the levels of fluid, and quality and human errors. These failures give rise to equipment breakdown and may cause some serious damage. According to an estimate, the breakdown of machines can be up to 20% [2].

Numerous studies have already been conducted by researchers on the fault classification of various systems including hydraulic systems. Adamsa et al. broke down the conventional problem into several subproblems under a hierarchical classification scheme to achieve maximum accuracy for Prognostic Health Monitoring (PHM). Furthermore, for each subproblem, reinforcement learning was suggested for the classifiers. They tested three reinforcement learning algorithms, i.e., Monte Carlo learning, Q-learning, and SARSA. The hydraulic actuator's condition was evaluated using the suggested methods [3].

Helwig et al. evaluated the classification of component conditioning using LDA by simulating fault situations such as valve changing deterioration, pump leaks, accumulator gas leaks, and oil dispersion, with the help of an experimental hydraulic test bench [4].

To tackle the issue of defect diagnosis utilizing bearing data given by the Case Western Reserve University (CWRU) data center, Yuan Y. et al. categorized the fault types using a CNN [5].

Nikolai et al. used supervised classification based on LDA for the evaluation of the statistical data acquired from the condition monitoring system. Automated feature extraction was used, and fault scenarios were diagnosed by using correlation criteria [6].

Cloud and edge servers were used by Fawwaz et al. to propose real-time fault recognition. Feature selection and offline learning were performed on the cloud, while online recognition close to the data source was computed at the edge. GA-based feature selection was used, and LSTM-AE was used as the fault detection model [7].

Keke Huang et al. discussed the problems faced in the acquisition of data in hydraulic systems such as multi-rate sensor data. To overcome the problem, a deep learning approach was proposed in their research work, which was able to automatically extract features from the data samples having multiple rates of acquisition. A CNN-based algorithm was designed to perform the classification of faults in hydraulic systems, which had a 10% better classification accuracy compared with the multiclass SVM [8].

Mallak et al. and Fathi et al. worked on the diagnostics and classification of hydraulic system faults. The proposed architecture was a combination of deep learning models, supervised machine learning models, and LSTM autoencoders. The system was tested on component faults, as well as sensor faults in a hydraulic system [9].

Silverstein et al. published research work on the comparison of different deep learning and traditional machine algorithms in the prediction of faults in hydraulic systems. They used the deep learning algorithms LSTM and TCN from, as well as machine learning algorithms: Decision Trees, k-Nearest Neighbors, and Random Forest. It was observed that TCN performed best, and enhanced feature engineering in traditional machine learning models also gave satisfactory results [10].

Yoo et al. presented an algorithm in their paper for the detection of faults using correlation-based clustering. The proposed approach was different from traditional clustering algorithms, as they decreased the size of the data to speed up the process of training the algorithm. The proposed approach clustered datasets of high correlation in a straight line; the distance of each dataset was calculated using the stochastic distance, and the abnormality detection index was also calculated. The proposed algorithm was applied to a dataset of a hydraulic system for verification [11].

Zhuo et al. and Z. Ge et al. published their research work in the field of industrial processes. The goal of the work was to develop a list of potential faults, with the power of existing data, that could occur in the future or have not been addressed before. The Generative Adversarial Network (GAN) was utilized to deal with this any-shot learning problem. The algorithm was tested on data acquired from a hydraulic system and TEP [12].

Konig et al. and Helmi et al. published a research paper in which they proposed a deep-learning-based model to provide data about the sensitivity and significance of the sensors in condition monitoring. Another objective of the research was to observe the contribution of each sensor to the overall result. Matthew's Correlation Coefficient (MCC) was used to determine the classifier's accuracy [13].

The research paper published by Kim et al. and Jeong et al. proposed a novel approach to fault classification in hydraulic systems. A setup was developed to track the condition of the overall system in real-time. Features were extracted from the input data using a CNN and a BiLSTM trained on the features; a sigmoid function was used as the classifier, and the data learned by BiLSTM were fed to the classifier. To overcome the problem of data scarcity, an augmentation technique was used to create new data. It was reported that the projected system performed better than the conventional deep learning approaches [14].

Yantao Zhu proposed a hybrid approach for dam deformation prediction by combining statistical regression, phase space reconstruction, and an improved LSTM neural network with parameter optimization using the GWO algorithm. The experimental results on a high-arch dam demonstrated effective noise elimination and high prediction accuracy, offering a balance between model performance and interpretability [15].

Leong et al., Ooi et al., and Lim et al. proposed an adaptive Genetic Algorithm to overcome problems faced in standard feature selection processes such as convergence towards local optima, manual parameter tuning, premature convergence, lower feature subset reduction rates, and the excessive cost of computation. The proposed STLA-GA was able to outperform classic feature-selection methods due to its adaptive nature [16].

A study conducted by Bo Jin applied deep transfer learning to detect the occurrence of diseases using facial diagnosis. The top 1 accuracy was greater than 90%, which was a better accuracy than the traditional machine learning method and clinical diagnosis [17].

Qinghe Zheng conducted a study for automatic modulation classification by implementing spectrum interference based on a two-level data augmentation method. After comparison with a variety of data augmentation techniques, it was concluded that the proposed method showed much advancement [18].

Nantian Huang proposed using a label-noise-robust Auxiliary Classifier Generative Adversarial Network (rAC-GAN) for fault diagnosis of rolling bearings for wind turbine gearboxes. The model demonstrated higher accuracy than other models available for multistate classification of rolling bearings [19].

Bin Cao studied how deploying sensor nodes and relay nodes in an industrial environment affected security, lifetime, and coverage issues. The author applied six serial algorithms and two parallel algorithms. This configuration allowed the author to achieve better performance in less time [20].

Hanxin Chen conducted a study in which the author proposed a diagnosis method to detect faults in a centrifugal pump using a three-dimensional matrix having time, frequency, and space as its dimensions. This study assisted in understanding the normal and faulty states of the mechanical equipment using improved particle swarm algorithms [21].

The study conducted by Yantao Zhu proposed an automatic damage-detection method for large-volume hydraulic concrete structures using drones and AI techniques. The approach combined computer vision with the Xception backbone network for crack feature extraction and an adaptive attention mechanism based on Deeplab V3+ for precise identification of cracks, achieving high-precision results with a 90.537% Intersection Over Union (IOU), 91.227% precision, 91.301% recall, and 91.264% F1-score [22].

Data augmentation was applied to the time series data of a hydraulic system to obtain a hefty amount of data for better classification because deep learning models do not work well when the amount of data is tiny. A five-layer Convolutional-Neural-Network-based Visual Geometry Group Network (VGGNet) was used for data augmentation [23].

Kortmann et al. presented a method for feature extraction; the approach utilized an unsupervised autoencoder for the task. The state of the cooler was set as the target variable from the dataset for classification purposes and regression, and the hydraulic accumulator

pressure was set. It was reported that the prediction accuracy was not as good as other approaches currently available for performing the same tasks [24].

To determine the condition of a hydraulic system, Y. Cheng et al. suggested a technique based on General multi-class Support Vector Machines (GenSVMs). The statistical features from the raw data were extracted during preprocessing such as the skewness, kurtosis, absolute mean value, and root mean square value. Four different GenSVMs were developed for four dissimilar components of the hydraulic system: pump, accumulator, cooler, and valve. The proposed model achieved a 100% classification accuracy for the pump, cooler, and valve and 76% for the accumulator. The proposed GenSVMs showed a better classification accuracy than the four other ML algorithms it was compared with [25].

ResNet-18 is a revolutionary deep convolutional neural network architecture introduced by Microsoft Research. It is differentiated from other methods through the incorporation of residual connections, also known as skip connections, which enable the learning of residual mappings and simplify the training of deeper networks. With 18 layers, including convolutional, pooling, and fully connected layers, ResNet-18 overcomes the vanishing gradient problem, making it easier for the network to learn effectively. The use of shortcut connections ensures that the input and output dimensions of each residual block remain the same, and additional  $1 \times 1$  convolutional layers are introduced when necessary to maintain consistency. Despite its depth, ResNet-18 is computationally efficient with fewer parameters. Pre-trained models on large datasets, such as ImageNet, are often utilized for transfer learning, making ResNet-18 a widely adopted and powerful model for various image-recognition tasks, with a significant impact on subsequent deep learning architectures. ResNet-18 is relatively deeper compared to earlier CNN architectures such as AlexNet or VGG-16. It has 18 layers, including convolutional layers, pooling layers, and fully connected layers, while traditional CNNs typically have fewer layers.

In this paper, the authors propose the prediction of the working behavior of cooling circuits in a hydraulic system. XGBoost and ReliefF were compared for the feature ranking technique. XGBoost is a library that implements machine learning algorithms under the Gradient Boosting framework, while ReliefF is an algorithm that takes a filter-method approach to feature selection to calculate feature scores to rank the top-scoring features. A DNN and ANN were used for the prediction of the cooling circuit [26].

The study aims to investigate techniques used for the categorization of Hydraulic Systems; furthermore, it explains a new technique for the classification of the fault conditions of hydraulic systems using an Artificial-Intelligence-based approach.

The dataset was prepared using the test rig. The rig comprised multiple sensors including pressure sensors, volume flow sensors, temperature sensors, and various sensors for monitoring motor power, cooling efficiency, vibration, cooling power, and system efficiency. It included failure scenarios that depict the fault state of four primary components. The data were analyzed using deep learning technology, comprising 43,680 attributes and 2205 instances. Thirty percent of the data were tested and validated, after which the data were trained on for classification. A hydraulic test rig was used to experimentally gather the dataset for this study. A primary working circuit and a secondary cooling–filtration circuit made up the test rig, and they were connected by an oil tank. The system measured process parameters such as pressures, volume flows, and temperatures, while cycling through constant load cycles that lasted 60 s. Throughout the testing, the state of four hydraulic components, i.e., the cooler, valve, pump, and accumulator, was quantitatively changed.

The primary working circuit was in charge of sending energy from the actuator to the pump. It was made up of an actuator, a valve, and a pump. The hydraulic fluid was cooled and filtered via the secondary cooling–filtration circuit. It was made up of a reservoir, a filter, and a cooler. The hydraulic fluid was kept in the oil tank. The fluid was also cooled and filtered as a result. Constant load cycles simulated the types of loading that a hydraulic system would encounter in a practical application. The process values measured included pressures, volume flows, and temperatures. These numbers were used to evaluate the hydraulic system's efficiency and the health of its parts.

The dataset contains 43,680 attributes and 2205 instances. The attributes are the features of the dataset, and they represent the values that were measured by the sensors on the hydraulic test rig. The instances represent a single constant load cycle of 60 s in various conditions.

The attributes in the dataset include the pressures, volume flows, temperatures, and the condition of four hydraulic components: the cooler, valve, pump, and accumulator. The faults in the dataset represent intentional changes that were made to the hydraulic system to simulate the different types of faults that could occur in a real-world system. These faults contributed to the process sensor measurements, which could be used to assess the performance of the hydraulic system and the condition of the hydraulic components.

The dataset is a valuable resource for researchers who are interested in studying the performance of hydraulic systems. The dataset can be used to train machine learning models that can be used to predict the performance of hydraulic systems under different operating conditions. The dataset can also be used to develop new fault detection and diagnosis methods for hydraulic systems. Other approaches have been applied by researchers to evaluate prediction accuracy for their study. In this study, ResNet-18 was used to evaluate the prediction accuracy. Hence, this study aimed to achieve higher prediction accuracy than other techniques available for deep learning. In Table 1, the prediction accuracies for different models are analyzed. This was used as a reference for the comparison of the prediction accuracy of ResNet-18.

**Table 1.** Different approaches applied on the hydraulic test rig data for fault classification and its accuracy.

Approach	Application	Prediction Accuracy	Researchers
PCA and XGBoost	Fault diagnosis for hydraulic directional valves	88%	Y. Lei et al. [9]
EGMSVM	Health evaluation of complex degradation systems	94.1%	Jun Wu et al. [10]
kNN and SVM	Fault diagnosis for hydraulic systems	96.7%	X. Zhao et al. [22]
Gen-SVM	Health estimation of hydraulic systems	94%	Y. Cheng et al. [27]

## 2. Materials and Methods

The hydraulic system dataset is open source and can be obtained from the University of California Irvine (UCI) machine learning repository [28]. The data were prepared using the test rig. The rig was comprised of multiple sensors including pressure sensors, volume flow sensors, temperature sensors, and various sensors for monitoring motor power, cooling efficiency, vibration, cooling power, and system efficiency. Constant load cycles of 60 s were repeated cyclically for measuring the process values, i.e., pressures, temperatures, and volume flows. Additionally, the dataset includes failure scenarios that depict the fault state of four primary components: cooler, valve, internal pump leakage, and accumulator.

The information provided by the UCL repository is in several text files. As shown in Figure 1, six distinct text files include the processed sensor readings and one file contains the intended fault circumstances.

The data were combined into a single data frame after being translated to comma-separated values. Using the filter option in Microsoft Excel, different files for the fault conditions of the cooler system of the hydraulic test rig were developed by extracting the data from the .csv file, as demonstrated in Figure 2.

A spectrogram is a visual representation of the signal strength over time at different frequencies. A spectrogram is created using the Fast Fourier Transform by passing through a digital process. The digitally sampled data are divided into blocks in the time domain, which are overlapping, and Fourier Transformed to determine the size of the frequency spectrum for each block. Then, for each block, a vertical line in the image is represented; it

is the midpoint of the block, which corresponds to the magnitude versus frequency for a specific moment. The time plots are then arranged side by side to create an image or three-dimensional surface and can be windowed in a variety of ways with a small amount of overlap. The computation of the signal's Short-Time Fourier Transform and this operation are related [29].

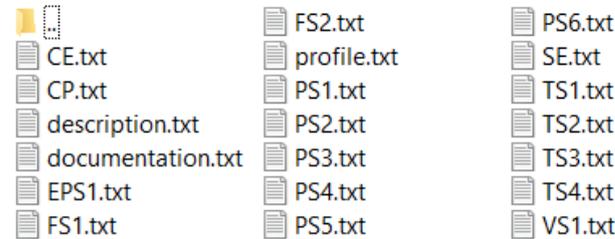


Figure 1. Different text files for processed sensor readings.

PS2	PS3	PS4	PS5	PS6	FS1	FS2	TS1	TS2	TS3	TS4	P1	VS1	CE1	CP1	SE1
109.467	1.991	0	9.842	9.728	6.71	10.305	35.622	40.979	38.471	31.745	2538.929	0.577	39.601	1.863	59.157
109.355	1.976	0	9.635	9.529	6.715	10.403	36.677	41.533	38.979	34.494	2531.499	0.566	25.786	1.256	59.336
109.159	1.972	0	9.531	9.428	6.719	10.366	37.881	42.442	39.632	35.646	2519.928	0.577	22.218	1.113	59.543
109.065	1.947	0	9.439	9.337	6.721	10.303	38.879	43.404	40.403	36.579	2511.542	0.569	20.46	1.062	59.795
108.931	1.923	0	9.359	9.261	6.69	10.238	39.804	44.333	41.311	37.428	2503.45	0.577	19.787	1.07	59.455
108.888	1.913	0	9.301	9.207	6.699	10.179	40.659	45.171	42.124	38.212	2501.007	0.573	19.15	1.072	59.563
108.676	1.888	0	9.234	9.143	6.699	10.141	41.464	45.947	42.888	38.932	2494.417	0.573	18.666	1.082	59.79
108.652	1.875	0	9.194	9.105	6.678	10.096	42.215	46.665	43.52	39.559	2489.422	0.572	18.178	1.079	59.591
108.53	1.858	0	9.145	9.057	6.672	10.04	42.892	47.339	44.201	40.081	2484.419	0.567	18.335	1.115	59.609
108.511	1.841	0	9.105	9.019	6.66	9.998	43.533	47.981	44.797	40.613	2480.435	0.572	18.206	1.129	59.474
108.438	1.832	0	9.068	8.984	6.659	9.964	44.106	48.586	45.365	41.1	2478.795	0.569	18.112	1.145	59.584
108.317	1.826	0	9.023	8.94	6.663	9.931	44.638	49.114	45.921	41.568	2472.273	0.578	18.069	1.165	59.758
108.241	1.816	0	8.981	8.902	6.656	9.914	45.16	49.602	46.391	42.038	2467.229	0.584	17.742	1.163	59.737
108.214	1.808	0	8.948	8.871	6.665	9.891	45.629	50.064	46.868	42.44	2465.307	0.589	17.678	1.18	59.916
108.127	1.804	0	8.915	8.841	6.669	9.857	46.059	50.509	47.318	42.799	2461.772	0.582	17.779	1.202	60.047
108.039	1.802	0	8.891	8.816	6.65	9.835	46.462	50.915	47.709	43.152	2459.511	0.582	17.692	1.207	59.811
107.981	1.794	0	8.89	8.813	6.65	9.799	46.84	51.299	48.063	43.446	2458.166	0.585	17.722	1.22	59.867
107.955	1.795	0	8.862	8.786	6.64	9.78	47.152	51.653	48.444	43.732	2455.588	0.593	17.82	1.242	59.768
107.837	1.793	0	8.837	8.763	6.66	9.761	47.47	51.996	48.74	44.044	2452.3	0.604	17.556	1.235	60.096
107.762	1.8	0	8.816	8.745	6.649	9.741	47.751	52.314	49.086	44.304	2448.326	0.603	17.664	1.256	60.019
107.794	1.8	0	8.82	8.751	6.638	9.696	48.026	52.585	49.347	44.498	2444.722	0.619	17.734	1.267	59.942
107.739	1.798	0	8.83	8.76	6.644	9.657	48.271	52.84	49.626	44.669	2441.776	0.631	17.943	1.291	60.187
107.73	1.791	0	8.833	8.764	6.637	9.632	48.5	53.062	49.856	44.855	2438.771	0.622	17.956	1.298	60.083
107.73	1.792	0	8.825	8.755	6.633	9.617	48.711	53.256	50.076	45.055	2437.651	0.625	17.93	1.302	59.959
107.712	1.793	0	8.814	8.743	6.635	9.607	48.92	53.462	50.269	45.208	2437.127	0.633	17.961	1.31	60.044
107.592	1.739	0	8.8	8.731	6.647	9.596	49.11	53.674	50.522	45.322	2434.606	0.636	18.289	1.345	60.351
107.517	1.731	0	8.796	8.727	6.635	9.574	49.319	53.871	50.721	45.496	2432.505	0.638	18.268	1.349	60.216
107.484	1.741	0	8.791	8.723	6.647	9.554	49.495	54.06	50.91	45.538	2432.652	0.627	18.646	1.383	60.431

Figure 2. Data compiled from different sensors' readings into a single .csv file for processing.

A continuous signal can be subjected to sampling to transform it into a discrete time signal by selecting values from the continuous time signal at evenly spaced points in time [27]. Hence, sampling a continuous-time signal  $x$  with sampling period  $T_s$  provides the discrete time signal  $x_s$ , defined by:

$$x_s(n) = x(nT_s) \quad (1)$$

Angular frequency sampling is given by:

$$\omega_s = 2\pi/T_s \quad (2)$$

A total of 440 spectrograms were generated for the targeted fault conditions of the cooler conditions: 146 spectrograms for Class Output 3 depicting the fault condition "close to failure" behavior of the cooler condition, 146 spectrograms for Class Output 20 depicting the fault condition "reduced efficiency" behavior of the cooler condition, and 148 spectrograms for Class Output 100 depicting the condition "full efficiency" behavior of the cooler condition.

AI by definition is the effort to automate intellectual tasks normally performed by humans (DL-w-P). It is a field of study that has enabled mankind to ponder how to integrate and utilize information, process and analyze data, and leverage the power of data to enhance the decision-making abilities of machines. The power of AI is being utilized in

a variety of sectors, such as finance, criminal justice, supply chains, automobiles, search engines, robotics, etc. [30].

ML is a major field of AI, and according to Tom Mitchell, “Machine learning is the study of computer algorithms that allow computer programs to automatically improve through experience” [31].

ML algorithms are majorly applied for regression, classification, clustering, recommender systems, and dimensionality reduction.

DL is regarded as an advancement of ML in which programmable NNs empower machines to make a judgment in the absence of human involvement. DL models are created such that they continually examine the data using a logical framework and behave similarly to humans’ decision-making abilities. In this regard, DL applications work on a complex structure of algorithms called Artificial Neural Networks (ANNs). The biological network of neurons in the human brain has influenced the design of ANNs. The learning systems of ANNs are considerably more competent than those of classical ML models. The DL model cannot make erroneous conclusions. However, similar to most AI cases, it also needs several trainings to ensure the learning process is true. However, when executed completely and properly, deep learning is shown to be a scientific marvel, and it is considered the backbone of true AI [32]. There are several DL algorithms used in DL models. Some of the widely used algorithms are CNNs, LSTM, RNNs, MLPs, SOMs, and autoencoders.

Residual Network-18 (ResNet-18) is a Convolutional Neural Network that is 18 layers deep.

The convolutional layer carries out the largest portion of the network’s computational burden. This layer receives the picture as the input and applies various convolutional procedures to it. It has several filters, sometimes known as kernels, and all of the training covers the parameters. In general, the filters are smaller than the original image. To create an activation map, the filters are convolved with each picture. Every component of the filter and image dot product is computed at every place as the filter advances over the picture’s width and height. This process is repeated for every element of the input layer to generate an activation map [33]. The main objective of the convolutional layer is the extraction of information and patterns from an image. Filters or kernels are in charge of retrieving low-level characteristics such as the color, gradient direction, etc., from the beginning of the networks. On the other hand, higher-level characteristics such as image edges are extracted by filters or kernels farther down the network.

The weighted sum of the input from a node or nodes is transformed into the output in a layer of the NN by such a function. The performance and capability of the neural network mainly depends on the choice of such a function. Different functions are needed for various model components. In most cases, networks are built so that each node in a layer uses the same function. However, such a function is used either following or during the internal processing of each network node.

A network normally consists of layers: an input layer in which raw data are fed from the domain, hidden layers that pass the output to an additional layer by taking the input from another layer, and output layers responsible for the predictions. Typically, the same activation function is used by all hidden layers. Different activation functions from the hidden layers are used by the output layer, and this depends on the type of prediction required by the model.

CNNs are specially designed for object detection or image processing. CNNs are normally used to empower computer vision by teaching the machine about processing the visual world. Facial Recognition Technology is one of the common uses of computer vision. To make it simple, a binary representation of the data is given as the digital picture. It has a collection of pixels that are organized in a grid-like pattern and carry data about each pixel, including its color and brightness.

The pooling layer receives the corrected feature map as the input. A series of these operations, max pooling to be precise, on an image is performed by the pooling layer. The dimensions of the feature map are reduced by the operation of down-sampling via pooling.

This reduces the amount of computation and weights needed in the network. The resultant two-dimensional array from the pooled feature map is flattened into a single, continuous, linear vector by the pooling layer [34]. To make the network more flexible, pooling merges many pixel values into a single one. This helps to decrease the chances of over-fitting, biasing networks towards particular pixels [35].

The Confusion Matrix technique is used to describe the summary of the performance measurement of a machine learning categorization. It gives a better idea about the classification model regarding whether making the right predictions and the measure of the errors it made. The performance of two or more classes can be measured by the amount of correct and faulty forecasts, a count value summary that is split into the classes.

It shows the confusion of the classification models while making predictions. It not only presents insights into the errors, but also the types of errors made by the classifier:

TP = number of True Positives;

FP = number of False Positives;

TN = number of True Negatives;

FN = number of False Negatives.

Following are the performance measures that can be observed by the Confusion Matrix.

The percentage of accurate forecasts with respect to all other forecasts is known as the accuracy. It is calculated using the formula:

$$\text{Accuracy}(\%) = \frac{(\text{TP} + \text{TN})}{(\text{TP} + \text{TN} + \text{FN} + \text{FP})} \times 100 \quad (3)$$

It evaluates the binary classification system by classifying the example into “positive” and “negative”. The F-score is a measure that combines the model’s recall and precision. It can be regarded as the “Harmonic Mean” of the model. It can be calculated using the formula:

$$\text{F - Score}(\%) = \left( \frac{2 \times (\text{Recall} \times \text{Precision})}{(\text{Recall} + \text{Precision})} \right) \times 100 \quad (4)$$

The recall is the proportion of correctly predicted positive outcomes to all positive outcomes. It can be calculated using the formula:

$$\text{Recall}(\%) = \frac{\text{TP}}{(\text{TP} + \text{FN})} \times 100 \quad (5)$$

The precision is the ratio of accurate positive forecasts to all positive predictions. It can be calculated using the formula:

$$\text{Precision}(\%) = \frac{\text{TP}}{(\text{TP} + \text{FP})} \times 100 \quad (6)$$

Residual Networks (ResNets) are extremely successful due to their relative improvement of around 28%, replacing the other architectures. They can be efficiently trained with 100 layers, as well as 1000 layers [36]. Stacking additional layers in Deep Neural Networks works efficiently with improved accuracy and performance for solving complex problems. The ResNet architecture was adopted for the fault classification of hydraulic systems.

The model was trained using MATLAB with 20 epochs for each iteration with the pool size being 3,3. Table 2 describes the parameters from the hydraulic test rig that were analyzed. Table 3 classifies the fault conditions of the hydraulic test rig depending on the class output.

The performance of a machine learning model is assessed using evaluation metrics. They are commonly employed for assessing a model’s performance over time or to compare different models.

The following are some of the most-popular evaluation metrics:

Accuracy: This is the proportion of predictions that are valid. Although it is the easiest metric to comprehend, it occasionally contains errors. For instance, even if a model is not

particularly effective at forecasting the minority class, it will have a high accuracy if it consistently forecasts the majority class.

**Table 2.** Parameters from the hydraulic test rig.

Pressure Sensor Data	Temperature Sensor Data	Volume Flow Data	Pump Efficiency, Cooling Efficiency, Vibration, and Efficiency Factor
PS1	TS1	FS1	EPS1
PS2	TS2	FS2	CE
PS3	TS3		CP
PS4	TS4		VS1
PS5			SE
PS6			

**Table 3.** Details of the fault condition of the hydraulic test rig depending on the class output.

Fault Type	Unit	No. of Classes	Class Output	Remark
Cooler condition	%	3	3	Close to total failure
			20	Reduced efficiency
			100	Full efficiency
Valve condition	%	4	73	Close to total failure
			80	Severe lag
			90	Small lag
			100	Optimal switch behavior
Accumulator	bar	4	90	Close to total failure
			100	Severely reduced pressure
			115	Slightly reduced pressure
			130	Optimal pressure
Internal pump leakage	-	3	0	No leakage
			1	Weak leakage
			2	Severe leakage

**Precision:** The percentage of accurate positive predictions is represented by this number. A model might predict that 100 out of 1000 cases will be positive, and if 90 of those predictions come true, then its precision is 90%.

**Recall:** This represents the percentage of real positives that were accurately anticipated. For instance, if the dataset contains 100 actual positives and the model accurately predicts 90 of them, the recall is 90%.

**F-score:** This represents the weighted average of the recall and precision. It is frequently used as a lone metric to sum up a model's performance.

The Confusion Matrix is a table that displays the quantity of true positives, false positives, false negatives, and true negatives in order to summarize the performance of a model. It can be used to determine the F-score, accuracy, precision, and recall.

It is impossible to exaggerate the significance of evaluation metrics. They are crucial for comprehending a machine learning model's performance and for making defensible choices regarding how to enhance the model.

Additional information on the significance of evaluation measures is provided below:

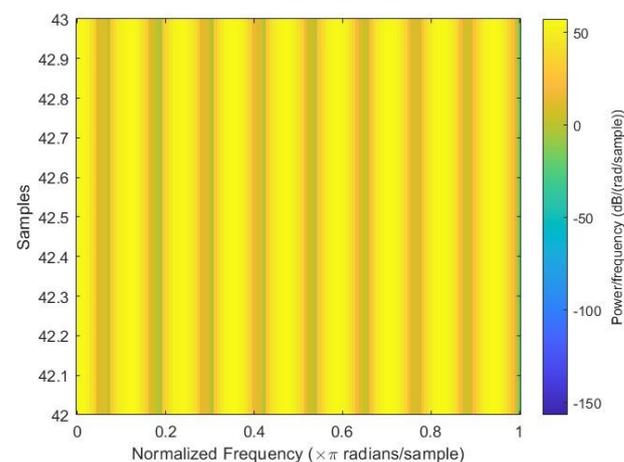
For various tasks, different metrics are appropriate. For activities where the cost of false positives and false negatives is comparable, accuracy, for instance, is an appropriate statistic. However, accuracy is a better indicator for activities when the cost of false positives is significantly larger than the cost of false negatives.

It is crucial to take into account not just the accuracy, but also the full Confusion Matrix. The Confusion Matrix can shed light on the kinds of mistakes the model is making. For instance, if a model has high accuracy, but low recall, it is probably underestimating the number of positive examples.

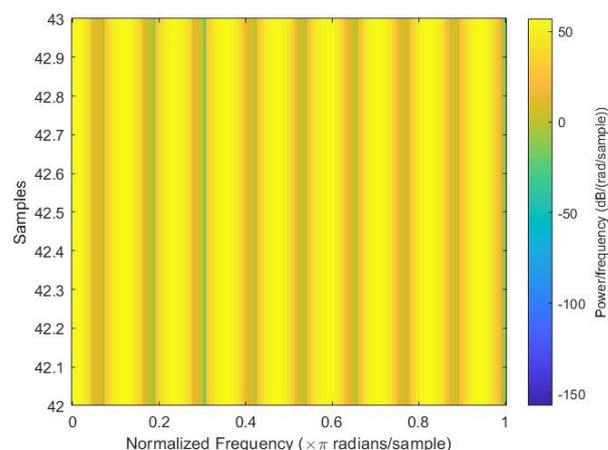
It is crucial to analyze a model using a variety of metrics. There is no one metric that can fully capture the performance of the model. One can have a better grasp of the model's advantages and disadvantages by using a variety of measurements.

### 3. Results

After all the spectrograms were generated, they were classified into three categories: close to failure, reduced efficiency, and full efficiency. Close to failure was assigned a class output of 3; similarly, reduced efficiency and full efficiency were assigned class outputs of 20 and 100, respectively. Among the spectrograms, 146 were Class Output 3, 148 were Class Output 20, and 146 were Class Output 100 out of a total of 440 spectrograms. Figures 3–5 show the sample spectrograms with operational, close to failure, and failed hydraulic cooling system.



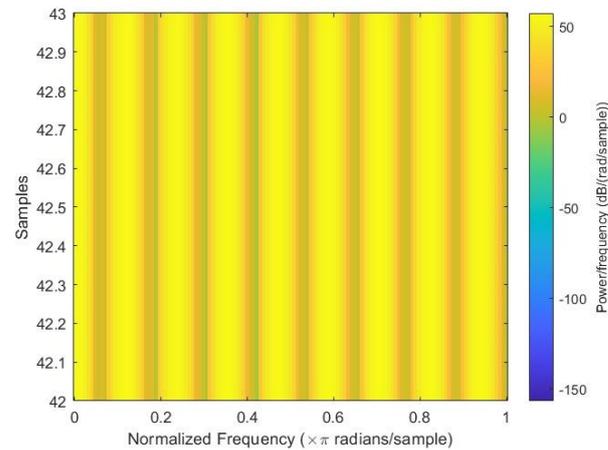
**Figure 3.** Sample spectrogram depicting the fault condition “close to failure” of the cooling circuit of the hydraulic test rig.



**Figure 4.** Sample spectrogram depicting the fault condition “reduced efficiency” of the cooling circuit of the hydraulic test rig.

The color of each point on the spectrogram represents the power of the radiation at that frequency and time. The spectrogram shows the power of the radiation from the hydraulic test rig. The different frequencies in the radiation can be attributed to the different components of the hydraulic system. The variation in the power of the radiation over time

can be attributed to the different fault conditions of the cooling circuit of the hydraulic system. Each figure has a slight difference from the others. In Figure 4, a sharp green line can be observed in the third power radiation; however, all other power radiations are identical. In Figure 5, the 3rd and 4th power radiations include green lines that are not as sharp as those in Figure 4; however, the rest of the power radiations are identical, which makes it different from the other figures.



**Figure 5.** Sample spectrogram depicting the condition “full efficiency” of the cooling circuit of the hydraulic Test rig.

The Confusion Matrices shown in Figures 6–8 were generated for the training, validation, and testing of the model of the cooling system for the hydraulic test rig, respectively.

		Confusion Matrix			
		cooler100	cooler20	cooler3	
Output Class	cooler100	100 32.7%	0 0.0%	0 0.0%	100% 0.0%
	cooler20	2 0.7%	99 32.4%	0 0.0%	98.0% 2.0%
	cooler3	0 0.0%	3 1.0%	102 33.3%	97.1% 2.9%
		98.0% 2.0%	97.1% 2.9%	100% 0.0%	98.4% 1.6%
		Target Class			
		cooler100	cooler20	cooler3	

**Figure 6.** Confusion Matrix of the training of the model of the fault condition of the cooling system of the hydraulic test rig.

In the training case, the overall accuracy of the model was found to be 98%. The model could correctly distinguish 98% of the fault condition “reduced efficiency” and 99% of both the “close to total failure” and “full efficiency” conditions of the cooling system of the hydraulic system. The results showed the high accuracy of the training of the model.

The overall precision of the training model was calculated as 98%. The precision for the fault condition “close to total failure” was 100%. The precision for the “reduced

efficiency” fault condition was 97% and for the “full efficiency” condition of the cooling system of the hydraulic system was 98%.

Output Class	Target Class			
	cooler100	cooler20	cooler3	
cooler100	21 31.3%	0 0.0%	0 0.0%	100% 0.0%
cooler20	0 0.0%	21 31.3%	1 1.5%	95.5% 4.5%
cooler3	2 3.0%	1 1.5%	21 31.3%	87.5% 12.5%
	91.3% 8.7%	95.5% 4.5%	95.5% 4.5%	94.0% 6.0%

**Figure 7.** Confusion Matrix of the validation of the model of the fault condition of the cooling system of the hydraulic test rig.

Output Class	Target Class			
	cooler100	cooler20	cooler3	
cooler100	22 32.8%	0 0.0%	2 3.0%	91.7% 8.3%
cooler20	0 0.0%	22 32.8%	0 0.0%	100% 0.0%
cooler3	1 1.5%	0 0.0%	20 29.9%	95.2% 4.8%
	95.7% 4.3%	100% 0.0%	90.9% 9.1%	95.5% 4.5%

**Figure 8.** Confusion Matrix of the testing of the model for the fault condition of the cooling system of the hydraulic test rig.

The overall recall of the training model was calculated as 98%. The recall for the condition “full efficiency” was 100%. The recall for the “reduced efficiency” fault condition was 98% and for the “close to total failure” fault condition of the cooling system of the hydraulic system was 97%.

The overall F-score of the training model was calculated as 98%. The F-score for the condition “full efficiency” was 99%. The F-score for the “reduced efficiency” fault condition was 97% and for the “close to total failure” fault condition of the cooling system of the hydraulic system was 98%.

For the validation case, the accuracy, precision, recall, and F-score were determined for different conditions of the cooling system of the hydraulic test rig.

The overall accuracy obtained for the validation case was 96%. The model could predict correctly 94% of the fault conditions for the close to total failure conditions. The accuracy for reduced efficiency and full efficiency was 97% for the cooling system of the hydraulic test rig.

The overall precision of the training model was calculated as 93%. The precision was 95%, 91%, and 95% for reduced efficiency, full efficiency, and close to failure, respectively.

The precision for the fault condition “full efficiency” was 91%. However, the precision for the “reduced efficiency” fault condition was 95% and for the “close to total failure” condition of the cooling system of the hydraulic system was 95%.

The overall recall of the training model was calculated as 94%. The recall was 95%, 87%, and 87% for the reduced efficiency, full efficiency, and close to failure, respectively.

The F-score of the training model was calculated as 93%. The F-score was 91%, 95%, and 95% for the conditions close to total failure, full efficiency, and reduced efficiency, respectively.

Similarly, in the case of testing, the accuracy, precision, recall, and F-score were calculated with the help of the Confusion Matrix.

For the testing, the overall accuracy obtained was 96%. The accuracy obtained for close to failure, reduced efficiency, and full efficiency was 95%, 100%, and 95%, respectively.

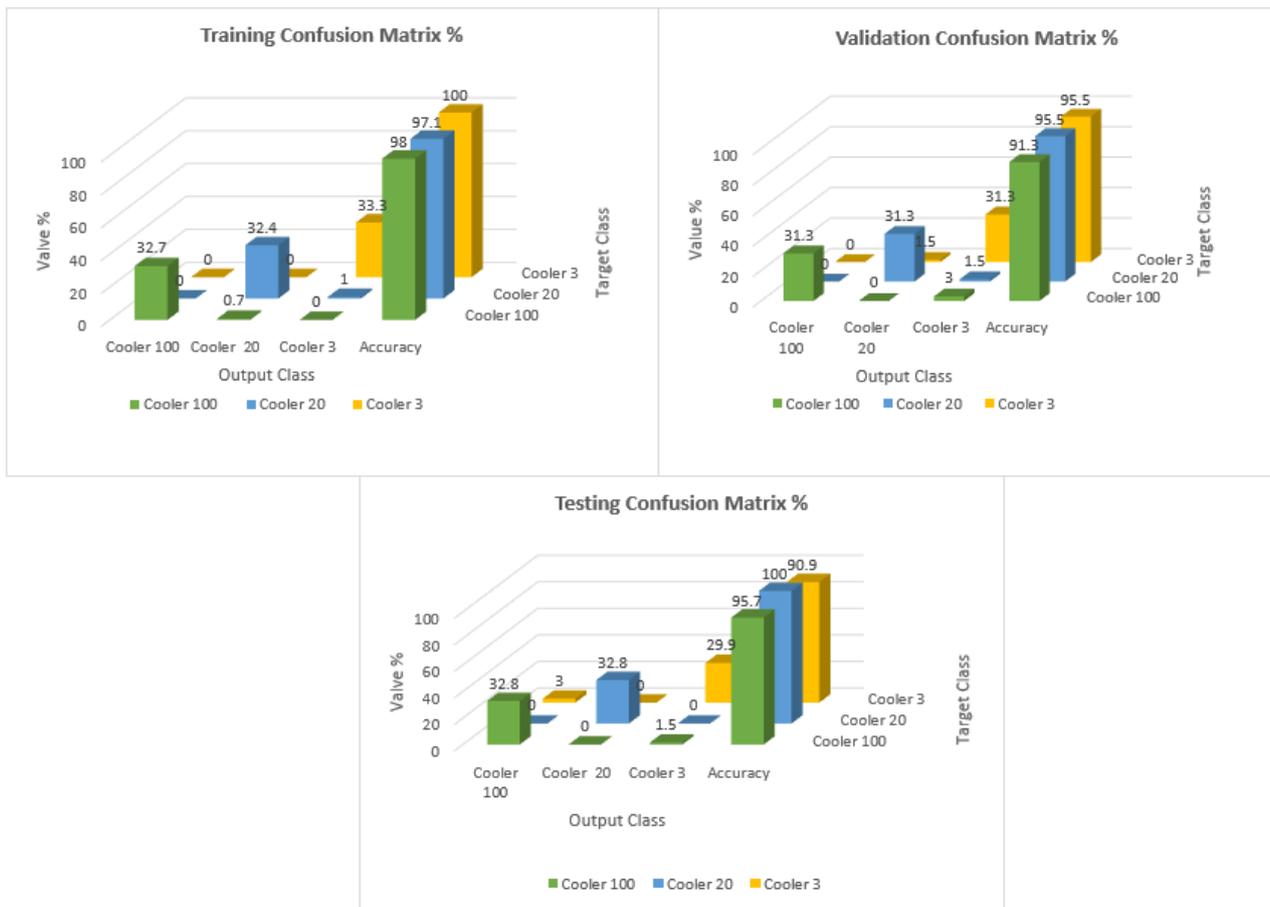
The overall precision obtained was 96%. The precision obtained for close to failure, reduced efficiency, and full efficiency was 91%, 100%, and 95%, respectively.

The total percentage recall for the testing obtained was 95%. The precision obtained for close to failure, reduced efficiency, and full efficiency was 91%, 100%, and 95%, respectively.

The overall F-score of the training model was calculated as 95%. The F-score obtained for close to failure, reduced efficiency, and full efficiency was 93%, 100%, and 93%, respectively. Figure 9 shows the distribution of the TP, TN, FP, and FN for each stage. Furthermore, Figure 10 shows the Confusion Matrix according to the class output for the training, testing, and validation.



Figure 9. Sample distribution for TP, TN, FP, and FN for training, testing, and validation.



**Figure 10.** Distribution and accuracy of the samples in the Confusion Matrix depending on the class output for training, testing, and validation.

#### 4. Discussion

After the literature review, different authors' models' accuracy were studied. Each author applied his/her technique for predicting fault conditions in the hydraulic test rig. However, in most cases, the model accuracy was not sufficient for predicting fault conditions.

The future aim of the research is to explore more suitable techniques and methods that could give greater accuracy. The impacts and results can be drawn by applying different techniques and comparisons of these results. The proposed technique can be applied for the classification of other faults of the hydraulic system, and the results can be drawn and compared. However, the classification of different sensors contributing to the condition monitoring of the hydraulic test rig can be used to conclude on the most-suitable and -reliable data of the sensors contributing to the result accuracy. ML techniques can be applied to the real-time data values of various systems for analysis. Real-time data can be processed, and more conclusions can be made. With time, more data can be extracted, and more-powerful predictions of fault conditions based on these data can be made, which can be a game changer in the Industrial Revolution.

#### 5. Conclusions

The data used in the thesis work were acquired from the UCI machine learning repository. The dataset includes failure scenarios that depict the fault state of four primary components of the hydraulic test rig: cooler, valve, pump leakage, and accumulator. The fault condition of the cooler was observed for classification in this thesis work. The data are in raw text form in different files. They were preprocessed to merge them into a single file. Spectrograms were then generated for each fault condition with the help of the data provided. These spectrograms were used in the Resnet-18 architecture of Convolution

Neural Networks for the classification. The results showed the high accuracy of the model, resulting in 96%, whereas the precision, recall, and F-score resulted in being 95%.

**Author Contributions:** Conceptualization, H.A.K., S.M. and K.K.; methodology, H.A.K., S.M. and U.B.; validation, K.K., M.A. and M.H.A.; formal analysis, H.A.K. and U.B.; investigation, M.A. and M.H.A.; resources, S.M., M.A. and M.H.A. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was supported by Researchers Supporting Project Number (RSP2023R274), King Saud University, Riyadh, Saudi Arabia.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** The data presented in this study are available upon request from the corresponding author.

**Acknowledgments:** This work was supported by Researchers Supporting Project Number (RSP2023R274), King Saud University, Riyadh, Saudi Arabia.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Hydraulic Equipment Market Size, Trends | Industry Forecast 2026. Available online: <https://www.technavio.com/report/hydraulic-equipment-market-industry-analysis> (accessed on 18 June 2023).
2. Rehorn, A.G.; Jiang, J.; Orban, P.E. State-of-the-art methods and results in tool condition monitoring: A review. *Int. J. Adv. Manuf. Technol.* **2005**, *26*, 693–710.
3. Adams, S.; Meekins, R.; Beling, P.A.; Farinholt, K.; Brown, N.; Polter, S.; Dong, Q. Hierarchical fault classification for resource constrained systems. *Mech. Syst. Signal Process.* **2019**, *134*, 106266. [CrossRef]
4. Helwig, N.; Klein, S.; Schütze, A. Identification and quantification of hydraulic system faults based on multivariate statistics using spectral vibration features. *Procedia Eng.* **2015**, *120*, 1225–1228.
5. Yuan, Y.; Ma, G.; Cheng, C.; Zhou, B.; Zhao, H.; Zhang, H.-T.; Ding, H. Artificial intelligent diagnosis and monitoring in manufacturing. *arXiv* **2018**, arXiv190102057.
6. Helwig, N.; Pignanelli, E.; Schütze, A. Condition monitoring of a complex hydraulic system using multivariate statistics. In Proceedings of the 2015 IEEE International Instrumentation and Measurement Technology Conference (I2MTC) Proceedings, Pisa, Italy, 11–14 May 2015; IEEE: Piscataway, NJ, USA, 2015; pp. 210–215.
7. Fawwaz, D.Z.; Chung, S.-H. Real-time and robust hydraulic system fault detection via edge computing. *Appl. Sci.* **2020**, *10*, 5933.
8. Huang, K.; Wu, S.; Li, F.; Yang, C.; Gui, W. Fault diagnosis of hydraulic systems based on deep learning model with multirate data samples. *IEEE Trans. Neural Netw. Learn. Syst.* **2021**, *33*, 6789–6801.
9. Mallak, A.; Fathi, M. Sensor and component fault detection and diagnosis for hydraulic machinery integrating LSTM autoencoder detector and diagnostic classifiers. *Sensors* **2021**, *21*, 433.
10. Silvestrin, L.P.; Hoogendoorn, M.; Koole, G. A Comparative Study of State-of-the-Art Machine Learning Algorithms for Predictive Maintenance. In Proceedings of the SSCI, Xiamen, China, 6–9 December 2019; IEEE: Piscataway, NJ, USA, 2019; pp. 760–767.
11. Yoo, Y. Data-driven fault detection process using correlation based clustering. *Comput. Ind.* **2020**, *122*, 103279.
12. Zhuo, Y.; Ge, Z. Auxiliary information-guided industrial data augmentation for any-shot fault learning and diagnosis. *IEEE Trans. Ind. Inform.* **2021**, *17*, 7535–7545.
13. König, C.; Helmi, A.M. Sensitivity analysis of sensors in a hydraulic condition monitoring system using CNN models. *Sensors* **2020**, *20*, 3307.
14. Kim, K.; Jeong, J. Real-time monitoring for hydraulic states based on convolutional bidirectional LSTM with attention mechanism. *Sensors* **2020**, *20*, 7099. [PubMed]
15. Zhu, Y.; Xie, M.; Zhang, K.; Li, Z. A Dam Deformation Residual Correction Method for High Arch Dams Using Phase Space Reconstruction and an Optimized Long Short-Term Memory Network. *Mathematics* **2023**, *11*, 2010. [CrossRef]
16. Ooi, C.S.; Lim, M.H.; Leong, M.S. Self-tune linear adaptive-genetic algorithm for feature selection. *IEEE Access* **2019**, *7*, 138211–138232.
17. Jin, B.; Cruz, L.; Gonçalves, N. Deep Facial Diagnosis: Deep Transfer Learning From Face Recognition to Facial Diagnosis. *IEEE Access* **2020**, *8*, 123649–123661. [CrossRef]
18. Zheng, Q.; Zhao, P.; Li, Y.; Wang, H.; Yang, Y. Spectrum interference-based two-level data augmentation method in deep learning for automatic modulation classification. *Neural Comput. Appl.* **2021**, *33*, 7723–7745. [CrossRef]
19. Huang, N.; Chen, Q.; Cai, G.; Xu, D.; Zhang, L.; Zhao, W. Fault Diagnosis of Bearing in Wind Turbine Gearbox Under Actual Operating Conditions Driven by Limited Data With Noise Labels. *IEEE Trans. Instrum. Meas.* **2021**, *70*, 3502510. [CrossRef]

20. Cao, B.; Zhao, J.; Gu, Y.; Fan, S.; Yang, P. Security-Aware Industrial Wireless Sensor Network Deployment Optimization. *IEEE Trans. Ind. Inform.* **2020**, *16*, 5309–5316. [[CrossRef](#)]
21. Chen, H.; Xiong, Y.; Li, S.; Song, Z.; Hu, Z.; Liu, F. Multi-Sensor Data Driven with PARAFAC-IPSO-PNN for Identification of Mechanical Nonstationary Multi-Fault Mode. *Machines* **2022**, *10*, 155. [[CrossRef](#)]
22. Zhu, Y.; Tang, H. Automatic Damage Detection and Diagnosis for Hydraulic Structures Using Drones and Artificial Intelligence Techniques. *Remote Sens.* **2023**, *15*, 615. [[CrossRef](#)]
23. Kim, K.; Jeong, J. Deep learning-based data augmentation for hydraulic condition monitoring system. *Procedia Comput. Sci.* **2020**, *175*, 20–27.
24. Kortmann, K.-P.; Fehsenfeld, M.; Wielitzka, M. Autoencoder-based Representation Learning from Heterogeneous Multivariate Time Series Data of Mechatronic Systems. *arXiv* **2021**, arXiv210402784.
25. Cheng, Y.; Zhu, H.; Wu, J.; Guo, P. Health State Estimation of Hydraulic System Based on Generalized Multiclass Support Vector Machine. In *Structural Health Monitoring 2019*. 2019. Available online: <https://www.dpi-proceedings.com/index.php/shm2019/article/view/32409> (accessed on 8 August 2023).
26. Prakash, J.; Kankar, P.K. Health prediction of hydraulic cooling circuit using deep neural network with ensemble feature ranking technique. *Measurement* **2020**, *151*, 107225.
27. 10.1: Signal Sampling. Available online: [https://eng.libretexts.org/Bookshelves/Electrical\\_Engineering/Signal\\_Processing\\_and\\_Modeling/Signals\\_and\\_Systems\\_\(Baraniuk\\_et\\_al.\)/10%3A\\_Sampling\\_and\\_Reconstruction/10.01%3A\\_Signal\\_Sampling](https://eng.libretexts.org/Bookshelves/Electrical_Engineering/Signal_Processing_and_Modeling/Signals_and_Systems_(Baraniuk_et_al.)/10%3A_Sampling_and_Reconstruction/10.01%3A_Signal_Sampling) (accessed on 18 June 2023).
28. Nikolai Helwig, E.P. Condition Monitoring of Hydraulic Systems. *UCI Mach. Learn. Repos.* **2018**. [[CrossRef](#)]
29. Smith, J.O. *Mathematics of the Discrete Fourier Transform (DFT): With Audio Applications*; W3K Publishing: Singapore, 2008; ISBN 0-9745607-4-X.
30. How Artificial Intelligence Is Transforming the World. Available online: <https://www.brookings.edu/research/how-artificial-intelligence-is-transforming-the-world/> (accessed on 18 June 2023).
31. Mitchell, T.M. *Machine Learning*; McGraw-Hill: New York, NY, USA, 2007; Volume 1.
32. Deep Learning, vs. Machine Learning: What's the Difference? Available online: <https://www.zendesk.com/blog/machine-learning-and-deep-learning/> (accessed on 18 June 2023).
33. Mostafa, S.; Wu, F.-X. Chapter 3—Diagnosis of Autism Spectrum Disorder with Convolutional Autoencoder and Structural MRI Images. In *Neural Engineering Techniques for Autism Spectrum Disorder*; El-Baz, A.S., Suri, J.S., Eds.; Academic Press: Cambridge, MA, USA, 2021; pp. 23–38, ISBN 978-0-12-822822-7.
34. Top 10 Deep Learning Algorithms You Should Know in 2023. Available online: <https://www.simplilearn.com/tutorials/deep-learning-tutorial/deep-learning-algorithm> (accessed on 18 June 2023).
35. Building A Convolutional Neural Network—The Click Reader. Available online: <https://www.theclickreader.com/building-a-convolutional-neural-network/> (accessed on 18 June 2023).
36. Team, G.L. Introduction to Resnet or Residual Network. Available online: <https://www.mygreatlearning.com/blog/resnet/> (accessed on 18 June 2023).

**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.