

Article

A Hybrid Fault Diagnosis Approach Using FEM Optimized Sensor Positioning and Machine Learning

Sang Jin Jung¹, Tanvir Alam Shifat²  and Jang-Wook Hur^{3,*}¹ LIG Nex1, Integrated Product Support R&D, 354-25, Sanho-daero, Gumi 39262, Korea² Department of Mechanical Engineering, Kumoh National Institute of Technology, Gumi 39177, Korea³ Department of Aeronautics, Mechanical and Electronic Convergence Engineering, Kumoh National Institute of Technology, Gumi 39177, Korea

* Correspondence: hhjw88@kumoh.ac.kr

Abstract: Sensor acquired signal has been a fundamental measure in rotary machinery condition monitoring (CM) to enhance system reliability and stability. Inappropriate sensor mounting can lead to loss of fault-related information and generate false alarms in industrial systems. To ensure reliable system operation, in this paper we investigate a system's multiple degrees-of-freedom (DOF) using the finite element method (FEM) to find the optimum sensor mounting position. An appropriate sensor position is obtained by the highest degree of deformation in FEM modal analysis. The effectiveness of the proper sensor mounting position was compared with two other sensor mounting points, which were selected arbitrarily. To validate the effectiveness of this method we considered a gear-actuator test bench, where the sensors were mounted in the same place as the FEM simulation. Vibration data were acquired through these sensors for different health states of the system and failure patterns were recognized using an artificial neural network (ANN) model. An ANN model shows that the optimum sensor mounting point found in FEM has the highest accuracy, compared to other mounting points. A hybrid CM framework, combining the physics-based and data-driven approaches, provides robust fault detection and identification analysis of the gear-actuator system.



Citation: Jung, S.J.; Shifat, T.A.; Hur, J.-W. A Hybrid Fault Diagnosis Approach Using FEM Optimized Sensor Positioning and Machine Learning. *Processes* **2022**, *10*, 1919. <https://doi.org/10.3390/pr10101919>

Academic Editors: Ming-Jong Tsai and Ricky Min-Fan Lee

Received: 25 August 2022

Accepted: 14 September 2022

Published: 22 September 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 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/>).

Keywords: condition monitoring; ANN; finite element method; deep learning; gear pump

1. Introduction

Prognostics and health management (PHM) for engineering systems has become a primary element to ensure safety and reliable operation. A robust PHM framework can improve reliability and eliminate the risk of catastrophic failures in a system. There are two major steps in PHM: fault diagnosis and fault prognosis. Finding an anomaly in a system's operation, or irregular behavior that is not the expected normal behavior, is treated as fault diagnosis. Fault prognosis largely depends on the remaining useful life (RUL) estimation of the component, based on its historical fault diagnosis data. Therefore, it is understandable that fault detection and identification (FDI) is the fundamental and most important task in a PHM framework [1]. Nowadays, the following three approaches to PHM are widely adopted: (a) Data-driven (b) Physics-based, and (c) Hybrid, combining (a) and (b). The major advantage of a data-driven PHM framework is that it does not require a prior mathematical model of the system to assess its health state [2]. Various data acquisition (DAQ) systems, data analysis tools, modeling using artificial intelligence techniques, etc., have made data-driven technologies popular among researchers. On the other hand, a physical model is the core of a physics-based PHM where the system material and structural properties are taken into consideration [3]. Data quality is an essential component of a data-driven PHM framework, since the state of health is determined based on this. It is also quite important in the physics-based approach to describe and validate the mathematical model [2,3]. Depending on the necessity of FDI nature, different sensors can be placed in different components of a system. Some commonly used sensor data for

FDI are temperature, current, voltage, impedance, vibration, torque, speed, magnetism, etc. In the case of rotary machinery, the most frequently used data are the vibrational data, which are acquired using accelerometer sensors.

A vibration signal can capture the fault characteristics at the incipient stage of a failure, and it can be easily mounted on a system to acquire a large amount of data [4]. Ease in data acquisition and modeling using different statistical and machine learning approaches have given the advantage to vibration signal monitoring for system condition monitoring. Several signal processing and model-based FDI models have been established using the vibration data acquired from sensors. However, there are very limited studies on the proper sensor positioning for accurate data acquisition. The authors of this paper took a keen interest in this fact, and a study was conducted on proper sensor positioning by combining a physics-based FEM simulation and data-driven FDI modeling. In literature, sensor positioning for system reliability enhancement has been studied to some extent. Krysnader et al. developed a sensor placement algorithm based on the structural information in a model [5]. Sarrate et al. proposed a strategy based on diagnosability maximization for optimally locating sensors in distribution networks [6]. Szytyber et al. proposed a graph-based method by incorporating faults into a model [7]. Finite element analysis or finite element modeling (FEM) has been studied in different electromechanical systems for fault detection and diagnosis. Most of these studies are based on a simulated fault characteristic and do not include results from an experimental prototype. Some FEM studies include the validation of experimental prototypes [8,9]. For example, Ahmed et al. used an FEM based calibration approach for piezoelectric impedance measurements with two empirical case studies [10]. Li et al. studied the induction motor's broken rotor bar fault diagnosis using time-stepping coupled FEM analysis [11]. Vaseghi et al. used FEM analysis for fault diagnosis and parameter identification in permanent magnet motors [12]. However, experimental research combining FEM and data-driven approaches has been very limited in the literature.

Deep learning (DL) algorithms have been quite popular among researchers, due to their ability to learn patterns from big data [13]. To establish a robust FDI model, the complete lifecycle data of a system is necessary, which is understandably a massive amount of data [14]. Traditional machine learning algorithms fail to provide high accuracy in the presence of big data. Several DL-based artificial intelligence (AI) techniques have been reported in the literature [14,15]. For example, attention-based neural networks and artificial neural networks (ANN) for electric motors [16,17], optimized self-organizing maps for bearing [18], convolutional neural nets for gearboxes [19], generative adversarial networks for rotary machinery [20], etc. Among them, ANN has been quite effective in machinery FDI since it can easily learn the fault patterns without any rigorous mathematical computations. Implementing ANN is also less time-consuming, compared to many other DL algorithms.

In this study, we incorporated FEM analysis with an ANN model to enable maximum accuracy in machinery FDI. Instead of only a physics-of-failure model, or only a sensor acquired data-driven model, we focused on the system dynamics in diverse operating conditions through the combination of both approaches. Given the current rise in demand, machines operate in a wide range of diverse applications making it a problematic task to design separate models for separate operating conditions. Relying on a hybrid PHM approach, combining the physical model parameters as well as sensor acquired data, can effectively detect and prominently isolate faults in a system [21]. To facilitate this, we needed to select a set of sensor and sensor mounting options to ensure the data quality and demonstrate system behavior. Improper sensor positioning and weak sensor data can lead to an erroneous health-state approximation. This is where a physics-based mesh analysis of the system was performed, using the finite element method (FEM). Based on the modal analysis and frequency components of random vibration found through fast Fourier transform (FFT) and power spectral density (PSD), an appropriate mounting point with the highest deformation was selected.

To validate the effectiveness of the proposed sensor positioning method, we considered a gear-actuator test rig setup. Gear pumps are energy conversion parts of hydraulic systems and have been widely used in aviation, construction machinery, and other industries. The gear pump converts the mechanical energy of the primary motion into liquid hydraulic energy and then transfers it to the entire hydraulic system. If the gear pump fails, the hydraulic system loses its function, often causing the entire system to shut down. Therefore, monitoring the health state of the gear pump is necessary. The sensor mounting approach presented in this paper enables the early FDI of gear-actuator setup in the industrial environment. The rest of the paper is structured as follows: Section 2 presents the proposed FDI method, Section 3 elaborates the theoretical review, Section 4 describes the method and experiments, and Section 5 provides the result analysis.

2. Proposed Method

A hybrid fault diagnosis framework, combining physics-based and data-driven approaches, is investigated. A physical mesh model is created using the FEM technique and based on random vibration deformation, and a suitable sensor mounting position is selected. Later, vibration signal fault features obtained from the FEM simulated position are modeled using a fully connected artificial neural network (ANN). This paper aims to improve accuracy in fault detection of gear pumps using sensor acquired vibration signals. A concise framework of the proposed method is shown in Figure 1. To have better understanding about the study's contribution, the proposed method is explained in individual sections below.

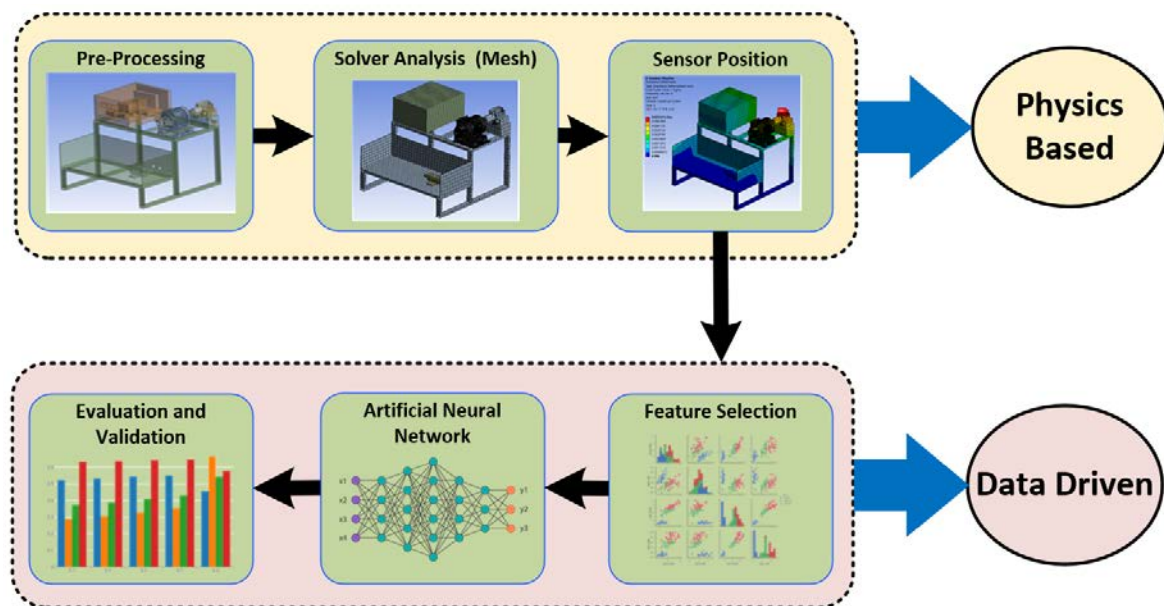


Figure 1. Proposed hybrid FDI method.

(1) FEM Analysis for Sensor Positioning

To determine a suitable sensor mounting location, a structural analysis of the experimental test bench was performed using the finite element method (FEM). In this FEM analysis, multi degrees-of-freedom (multi-DOF) dynamics of the experimental test rig were considered to perform the numerical analysis. FEM considers a multi-DOF system as a combination of some single-DOF systems by particular space discretization, which is constructed through the mesh layout of the experimental test rig.

(2) Selection of Weighty Features

Rotary machinery vibration signals are necessarily random in nature and hold variable fault characteristics in different health states. Analyzing the sensor acquired time series

data is not sufficient to conclude the state of health (SOH) of the system. Therefore, a range of statistical and rotary machinery features were extracted from the time domain, and frequency domain of the sensor acquired vibration signals.

Moreover, to select the most suitable features, we used three different criteria to ensure the elimination of redundant features and the selection of crucial features. The presence of less important features increases the computational time, and can sometimes lead to an erroneous approximation of the system's health state. This study proposed a two-stage analysis to select the best features. In the first stage, redundant features were eliminated, based on the similarity among features. Most similar or highly correlated features were excluded from consideration. Correlation analysis was performed by computing the Pearson correlation coefficient [22], which can be mathematically expressed using (1):

$$\rho = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}} \quad (1)$$

For the second stage of feature selection, a sensitivity index σ was computed using two different approaches, ANOVA and KW test score [23]:

$$\sigma = \frac{F_{score} + KW_{score}}{2} \quad (2)$$

where:

F_{score} = The ratio of the variance calculated among the means to the variance within the features, which is computed using the analysis of variance (ANOVA) test.

KW_{score} = Measures the stochastic dominance of each to one another.

(3) Pattern Recognition and Validation

The last stage was the classification of selected features using a deep neural network (DNN). There were two main reasons for selecting a neural network (NN) model for fault pattern recognition. Firstly, it is capable of handling big data. Secondly, it provides very high accuracy in future predictions. We used an artificial neural network (ANN) architecture to model the health states of the gear-actuator system. ANN is simple and does not include rigorous mathematical computations, unlike other NN models. From Step 1, two more additional sensor positions were selected to compare the effectiveness of the most suitable sensor. Steps 2 and 3 were repeated for the data acquired from all three sensors' positions. To validate the model performance, several matrices were computed from the training and testing performance of the models.

3. Theoretical Overview

3.1. Finite Element Method (FEM)

Finite element analysis (FEA) is a computer simulation technique used in engineering analysis. A numerical technique called the finite element method (FEM) is used, and, for its application, an object or system (system) must be expressed as a model similar to the real thing, as a plurality of individualized and interconnected finite elements. Equivalence equations are applied to each element, resulting in simultaneous equations in the overall system. An approximation can be obtained by solving an equation using linear algebra or nonlinear numerical analysis methods. As the method of obtaining an approximation is used, the accuracy of the FEA result can be improved through mesh refinement. A common use of FEA is in determining stresses and displacements in mechanical objects or systems. However, it can also be used for heat transfer, fluid mechanics, and electromagnetics. Modal analysis determines the dynamic properties (damping, natural frequency, and vibration mode) of a mechanical object or system. In time-position space, the system is independent of external forces, as well as dynamic responses. Systems with multiple degrees of freedom (multi-DOFs) can be specified and described by a group of combined systems of single-degree-of-freedom (single DOF). In that case, each scalar component, such as mass, stiffness,

and damping, is transformed into a matrix of mass, stiffness, and damping. The equation of motion describing a system with n degrees of freedom is expressed as [18]:

$$Mu''(t) + Cu'(t) + Ku(t) = f(t) \quad (3)$$

where M , C , and K are matrices of mass, damping, and stiffness, respectively. $u(t)$ and $f(t)$ are vectors that describe the change in displacement and force as a function of time, respectively. In addition, since the value of $u'(t)$ in (3) is always positive, we can intuitively see that this system is stable. However, using the Lyapunov stability theorem can mathematically prove the stability. Lyapunov function candidate is set using the following:

$$u'(t) = \frac{1}{2}Mu^2(t) + \int Ku(t) \quad (4)$$

$$u(t) = Mu''u' + Ku''u' = u'(Mu'' + Ku'') \quad (5)$$

As mentioned earlier, since $u'(t) > 0$, the Lyapunov function candidate is always negative except for $u'(t) \neq 0$, and if $u'(t) = 0$ equally, $u(t) \neq 0$. In the end, it can be seen that this system is a stable system because the time derivative of the Lyapunov function candidate is always negative under the condition that $u(t) \neq 0$, $u'(t) \neq 0$.

In general, if the system does not have an external force (i.e., $f(t) = 0$) and the damping is negligible (i.e., $C = 0$), Equation (3) can be modified and written as (6):

$$Mu''(t) + Ku(t) = 0 \quad (6)$$

In simple harmonic vibration, the position has the same relationship as in Equation (6).

$$u(t) = \{\varnothing_i\} \cos(\omega t) \quad (7)$$

where \varnothing_i is the amplitude vector and ω is the frequency of each degree of freedom i . The velocity is calculated as in (6), and the acceleration is calculated as in (7) [24,25].

$$u'(t) = -\omega\{\varnothing_i\} \sin(\omega t) \quad (8)$$

$$u''(t) = -\omega^2\{\varnothing_i\} \cos(\omega t) \quad (9)$$

If the velocity (refer to (8)) and acceleration (refer to (9)) are calculated by substituting in (7), we can arrive at the eigenvalue problem as shown in (10), which is:

$$(-M\omega^2 + K)\{\varnothing_i\} = 0 \quad (10)$$

In (9), since the amplitude $\varnothing_i \neq 0$, we can find the solution to the equation.

$$\det[-M\omega^2 + K] = 0 \quad (11)$$

In (11), we need to solve the eigenvalue problem and determine the n th order characteristic polynomial. The eigenvalue ω_i is equal to (11) and represents natural frequencies. Each eigenvalue ω_i is related to the amplitude \varnothing_i representing the vibration mode. For any geometry, $\Omega_i = \Omega_1 \times \Omega_2 \times \dots \times \Omega_n$. Equations (10) and (11) are obtained for the subdomain using the finite element methodology [26]. It is important to note that the mass and stiffness properties are fundamental factors in determining the natural frequency values. In this study, the mass parameters of the gear-actuator system were determined from the material properties, and the stiffness (dependence on elastic behavior) was obtained through sensitivity analysis using the experimental vibration data. Random vibration analysis was done through the PSD values obtained from the acquired vibration signals. Thus, based on the maximum mode deformation, a proper mounting position was selected. A concise FEM technique is shown in Figure 2.

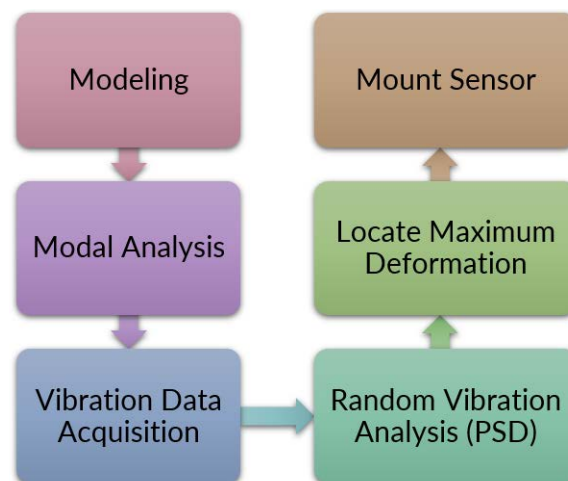


Figure 2. FEM architecture.

3.2. Pattern Recognition

Deep learning (DL) algorithms have proven quite effective for machinery FDI. Sensor acquired big data facilitates fault-related information in different system health states and allows deep learning algorithms to learn fault patterns efficiently. The nature and quality of the data govern the effectiveness of deep learning algorithms. A major step of DL modeling is data preprocessing, which starts with feature selection. Based on the need for individual tasks, there are numerous methods to select necessary features for a DL model [27]. Since our pattern recognition model is problem-specific, we chose our criteria to select the most important features. A set of features with the highest score was determined. Section 2 provides a detailed explanation of the feature selection and sensitivity score computation in paragraph 2.

Despite the selection of significant features, the gear pump vibration dataset was large due to the complete lifecycle data. Each feature is represented as a tensor in the DL algorithm and exhibits a matrix with some features and a number of timesteps [27,28]. Having many dimensions in the feature space surges the volume of the entire sample space. This is computationally expensive and time-consuming. In the case of DL modeling, a large number of dimensions also creates other problems. Too many dimensions make the feature space sparser and sparser [29,30].

The DL algorithm quickly finds a “local” solution in a sparse dimension instead of a “global” generalization. This causes the DL algorithm to miss out on much training data, and, in the case of testing data, it fails to predict accurately. This phenomenon is known as “overfitting” in machine learning algorithms. To overcome this issue, we used several techniques in our dataset for dimensionality reduction. These were principal component analysis (PCA), linear discriminate analysis (LDA), t-stochastic neighbor embedding (TSNE), and random trees ensemble (RTE). It is difficult to judge the effectiveness of these algorithms simply by visualization. Therefore, in this study, we compared the model’s output for each dimension reduction algorithm. When the feature space was reduced to a dimension, an artificial neural network (ANN) was used to model different health states of the gear pump. Figure 3 shows that the ANN architecture used for this study was necessarily a feed-forward neural network with fully connected layers and backpropagation. Hypothetical computation of each layer is done by (12) for the 1st layer [31,32]. In the case of multiple layers, the activation computation is given as (13). Here, x_i is the i th input variable or selected features for gear pump diagnosis. We expected the ANN model to predict the gear pump’s health state based on the input variables x_i . Therefore, a hypothesis function is introduced as h_θ , that maps the input variables with different health states. θ_i are model parameters or weights of ANN model which are selected in such a way that

$h_{\theta}(x)$ is close to y for the specific training examples. As the hypothesis function tries to map the prediction to the training set, this type of learning is called supervised learning.

$$h_{\theta}(x) = \theta_0 + \theta_1 x \quad (12)$$

$$h_{\theta}(x) = \theta_0 + \theta_1 x_1 + \theta_2 x_2 + \cdots + \theta_n x_n \quad (13)$$

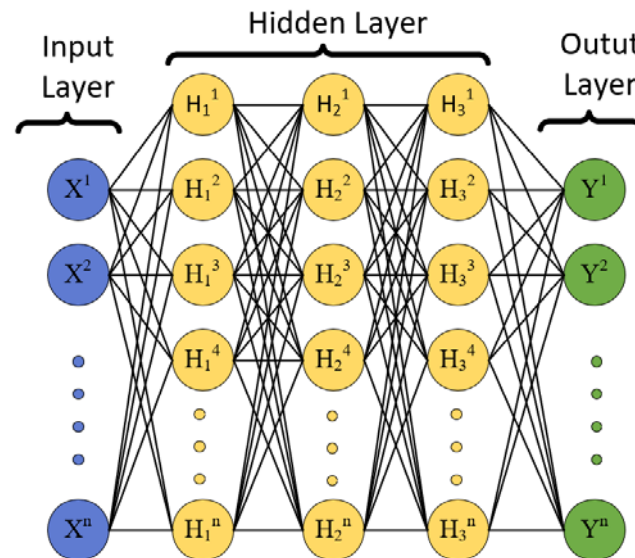


Figure 3. A simple ANN architecture.

These are hypothetical equations, and during training, each of these equations is governed by some activation function. In our model, we used rectified linear unit (ReLU) for the hidden layers and Softmax for the output layer. In ANN, a backpropagation algorithm is used to reduce the loss of the model, which is the difference between actual and predicted values. The loss function is mathematically shown in (14), which is also referred to as squared error function. The convergence of the loss function is determined by the computation of a gradient. The value m is referred to as the number of input variables or features. Gradient descent algorithm for multiple layers is shown in (15). Through this equation, θ_i, θ_j is simultaneously updated for every $j = 0, 1, 2, 3, \dots, n$.

$$J(\theta_0, \theta_1) = \frac{1}{2m} \sum_{i=1}^m (h_{\theta}(x^{(i)}) - y^{(i)})^2 \quad (14)$$

$$\theta_j := \theta_j - \alpha \frac{1}{m} \sum_{i=1}^m (h_{\theta}(x^{(i)}) - y^{(i)}) x_j^{(i)} \quad (15)$$

Based on the number of hidden layers, at each layer, the weight computation is done by following (16). Where $a_i^{(j)}$ = activation of unit i in layer j using the function $g(\cdot)$, and $\Theta^{(j)}$ is the linear combination of features presented beforehand.

$$a_i^{(j)} = g(\Theta^{(j)}) = g(\theta_{i0}^{(j-1)} x_0 + \theta_{i1}^{(j-1)} x_1 + \theta_{i2}^{(j-1)} x_2 + \cdots + \theta_{in}^{(j-1)} x_n) \quad (16)$$

Depending on the convergence time, computational cost, and prediction accuracy, model parameters, such as the number of hidden layers, activation function etc. were selected.

4. Experiment and Data Description

4.1. Test Rig Setup

The configuration of the experimental test bench is shown in Figure 4. The main components of the test rig were a hydraulic tank containing hydraulic oil, a gear pump to control hydraulic energy, a motor for driving the gear pump, a relief valve for generating hydraulic pressure by controlling the flow rate, a flow meter for measuring the flow rate, and a hydraulic meter for measuring the hydraulic pressure. Three different accelerometers were mounted at three different locations for measuring vibration signals. Vibration sensors were named Sensor A, Sensor B, and Sensor C, based on the FEM analysis results where A is the most appropriate position. All the accelerometer sensors were of a piezo electronic type and bought from the same manufacturer to reduce the discrepancy in sensitivity. Since vibration signals are random in nature and changes are abrupt in the presence of a fault or anomaly, we selected a 25.6 kHz sampling rate to acquire the vibration data. This allowed us to capture the frequency behavior for the gear pump for up to 12.8 kHz frequency band, as per the Nyquist theorem.

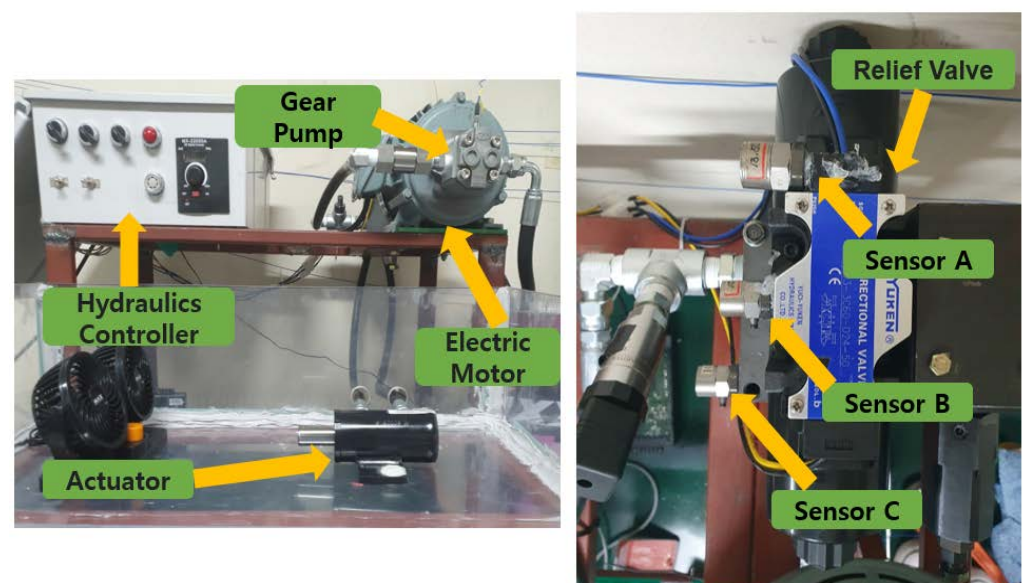


Figure 4. Test Bench Setup.

A data acquisition chassis, NI-cDAQ 9178, was used with the NI-9234 vibration module to acquire vibration data. An NI-9246 current module was also used to acquire the pressure and flow rate signals as a 0–20 mA current output. The mechanical properties of the materials for the analysis are shown in Table 1.

Table 1. Mechanical properties of the material.

Type	Plate	Bracket
Density	2700 kg/m ³	2830 kg/m ³
Young's Modulus	68.9 GPa	71.7 GPa
Poisson's Ratio	0.33	0.33
Shear Modulus	25.9 GPa	27.0 GPa
Yield Strength	275 MPa	490 MPa

When the motor is turned on, the gear pump is driven, and the gear pump draws hydraulic oil from the hydraulic tank and then compresses and discharges it. The discharged flow rate passes through the flow meter and the hydraulic system generates high pressure at the relief valve and returns the hydraulic oil to the hydraulic tank. During this process, the data acquired from each sensor is transmitted and stored to the computer through DAQ.

In the case of accelerated life tests, 1.0 g of SiO_2 was added to the hydraulic fuel every day to create disturbances in gear pump flow. Based on the flow rate and pressure reading, three different health states were selected from the gear pump manufacturer's provided datasheet. A healthy state (H) was considered to be when the gear pump's pressure and the flow rate were 80–100%. When the pressure dropped to 40–79% and the flow rate dropped to 50–79%, that state was labeled as Fault-1 (F1). At the severe failure stage, pressure dropped down to 0–39% and flow rate to 0–49%. This condition of operation was labeled as Fault-2 state (F2).

4.2. Sensor Positioning

Initially, three different locations were arbitrarily selected for triaxial (X-Y-Z) vibration data acquisition. The gear-actuator test rig mesh was analyzed using the output value of the modal analysis and the ASD data obtained through the random vibration FFT and PSD. Modal analysis was used to calculate the mode frequency and mode shape of the gear-actuator system, and random vibration analysis was applied to obtain the vibration characteristics under the load added by the input ASD data.

The FEM simulation results of the mode extraction and random vibration deformation are shown in Figure 5. The test rig mesh was divided into 395,400 nodes and 87,763 elements by applying a tetrahedral element and a hexahedral element. The frame, electric box, gear pump, and relief valve were attached in a bonded spherical contact condition, and the bottom of the frame was fixed with fixed constraints.

Sampling frequency for the random vibration signals was set to be $F_s = 25.6$ kHz. As per the Nyquist-Shannon theorem, effective observable mode frequency would be $F_s/2 = 12.8$ kHz. The modal analysis of the gear pump was performed until the mode up to 12.8 kHz was extracted to reflect all ASD data, and, at this time, a total of 154 modes were extracted, as shown in Table 2, and the 1st, 2nd, and 3rd modes were extracted. The shape of the mode is shown in Figure 5. It can be seen that the maximum displacement in the 1st mode occurred at the top of the relief valve, the maximum displacement in the 2nd mode occurred at the gear pump, and the maximum displacement in the 3rd mode occurred at the top of the electric box. Through this, it can be observed that the position of the maximum displacement changed according to the frequency.

The maximum equivalent stress and occurrence location, according to the gear pump analysis, are shown in Figure 5. The maximum equivalent stress was 41.9 MPa in the bracket in the vertical direction, 6.5 MPa in the aluminum structure in the transverse direction, and 0.9 MPa in the bracket in the longitudinal direction.

Therefore, the optimal position of the sensor was selected as the top of the relief valve, and data was collected by attaching two more sensors 5 cm and 10 cm away from the optimal position. The sensor mounted at the optimal position was labeled as Sensor A. The other two sensors were labeled as Sensor B and Sensor C. All three sensors kept collecting the vibration signals simultaneously during the entire lifecycle of the gear pump. A single data acquisition model was programmed using the LabVIEW software to maintain congruence among the sensors; for example, sampling rate, excitation current, starting and stopping time, sensitivity, etc., were set equal for all the sensors.

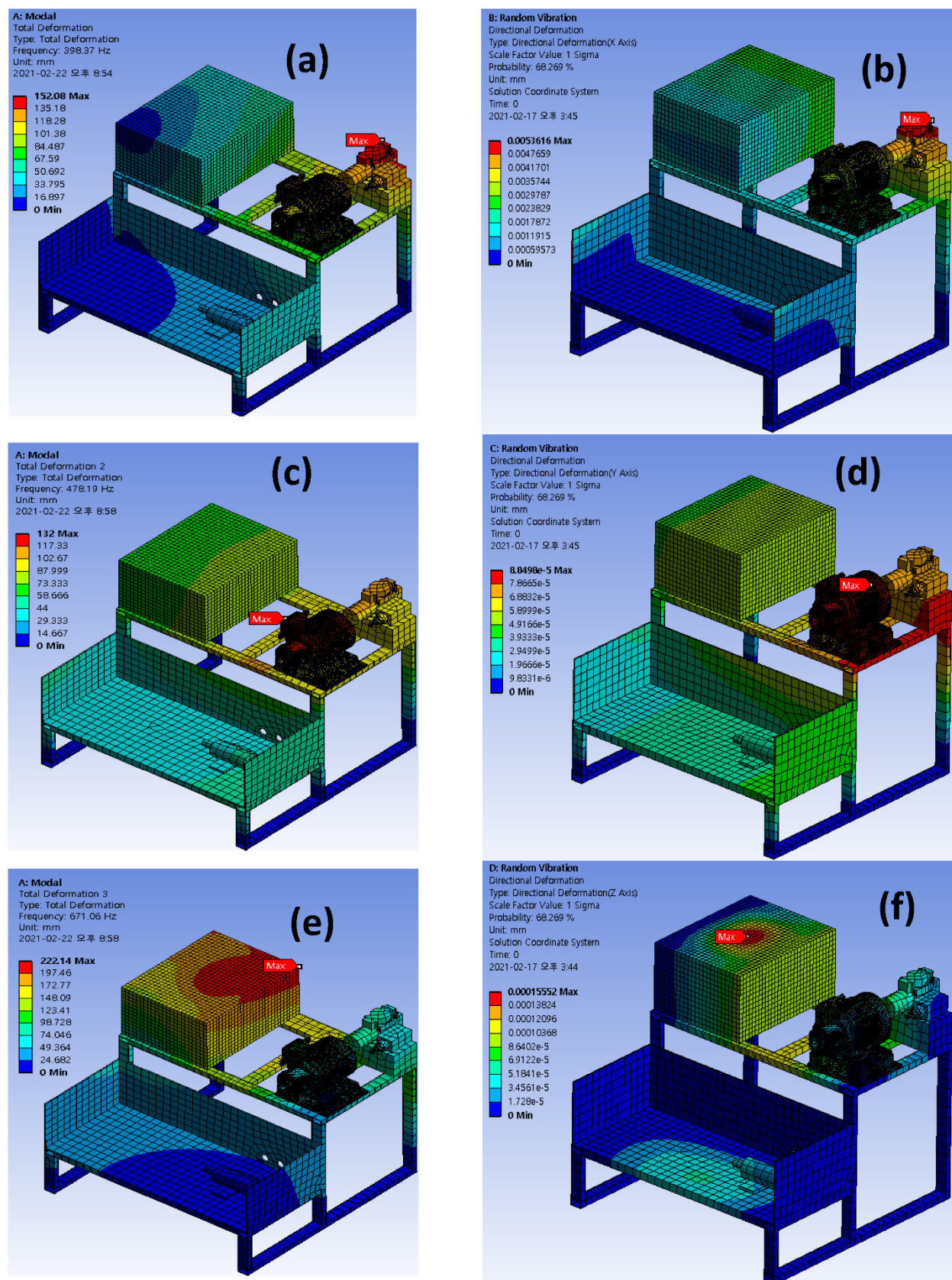


Figure 5. Mode extraction using FEM analysis. (a) 1st mode, (b) 2nd mode, (c) 3rd mode, (d) X-Axis, (e) Y-Axis, (f) Z-Axis.

Table 2. Mode extraction.

Mode (X)	Frequency (Hz)	Modal Mass (%)	Mode (Y)	Frequency (Hz)	Modal Mass (%)	Mode (Z)	Frequency (Hz)	Modal Mass (%)
1	578.829	45.78	1	578.829	9.32	1	578.829	11.43
2	732.294	2.09	2	732.294	68.40	2	732.294	4.37
3	761.892	30.06	3	761.892	8.01	3	761.892	21.48
4	1156.5	4.32	4	1156.5	7.69	4	1156.5	2.72
5	1338.29	8.06	5	1338.29	0.37	5	1338.29	0.11
6	2064.31	1.60	6	2064.31	0.15	6	2064.31	8.83
7	2133.15	1.88	7	2133.15	0.01	7	2133.15	1.42
8	2584.81	1.10	8	2584.81	0.00 *	8	2584.81	0.42
9	2646.56	0.07	9	2646.56	0.10	9	2646.56	0.14
10	2942.73	0.00 *	10	2942.73	2.05	10	2942.73	1.05
11	3134.17	0.02	11	3134.17	0.03	11	3134.17	3.38
12	3295.8	0.01	12	3295.8	1.01	12	3295.8	0.12
13	3597.87	0.12	13	3597.87	0.16	13	3597.87	6.96
14	3843.53	0.42	14	3843.53	0.17	14	3843.53	18.63
15	4053.49	0.88	15	4053.49	0.00 *	15	4053.49	5.73
16	4379.57	0.24	16	4379.57	0.01	16	4379.57	0.60
17	4638.19	0.13	17	4638.19	0.03	17	4638.19	1.92
18	4861.53	0.26	18	4861.53	0.01	18	4861.53	0.48
19	4975.44	0.17	19	4975.44	0.05	19	4975.44	1.29
20	5746.16	0.10	20	5746.16	0.00 *	20	5746.16	0.88

0.00 *: Indicates that it has a value after 4 decimal places.

5. Result Analysis

5.1. Data Analysis

To get a better understanding about the sensor signal, a range of various features was computed from all the sensor data [33]. A list of features computed from different domains is presented in Table 3 [33,34]. The most significant features were computed through a two-stage elimination and selection process. In the process of elimination, correlation among the features was computed using Pearson correlation coefficient, ρ . Similar features with $\rho \geq 0.8$ were eliminated. As mentioned in Section 1, we proposed a sensitivity index named σ , to select the most significant features. Based on the sensitivity score, $\sigma \geq 500$ were selected for further processing. Figure 6 presents the features, based on their σ score, and a brief description of the most significant features is presented in Table 4.

Table 3. The features computed in different dimensions.

Domains	Feature Names
Time Domain	Mean, Peak-to-Peak (P2P), Root Mean Square (RMS), Root Sum of Squares (RSSQ), Standard Deviation (STD), Kurtosis (KUR), Skewness (SKEW), L ¹ Norm (L1), L ² Norm (L2), Peak to RMS (P2RMS), Crest Factor (CF), Shape Factor (SF), Margin Factor (MF), Clearance Factor (CLF), FM4, FM8, M6A.
Frequency Domain	Peak Frequency (PF), Total Harmonic Distortion (THD), Spectral Skewness (SS), Spectral Kurtosis (SK), Entropy, Root Variance Frequency (RVF), SNR.

Table 4. Mathematical representation of selected features.

Label	Feature Name	Mathematical Expression	Label	Feature Name	Mathematical Expression
F1	Kurtosis	$\frac{N \sum_{n=1}^N [x(n) - \bar{x}]^4}{[\sum_{n=1}^N [x(n) - \bar{x}]^2]^2}$	F5	L2-Norm	$\left[\sum_i x_i ^2 \right]^{1/2}$
F2	RMS	$\sqrt{\frac{1}{N} \sum_{n=1}^N [x(n)]^2}$	F6	Root Variance Frequency	$\sqrt{\frac{\sum_{n=1}^N (f_n - f_{char})^2 x(n)}{\sum_{n=1}^N x(n)}}$
F3	Root Sum of Squares	$\sqrt{\sum_{n=1}^N [x(n)]^2}$	F7	Entropy	$-\sum_{n=1}^N \frac{x(n)}{\sum_{k=1}^N x_k} \log \frac{x(n)}{\sum_{k=1}^N x_k}$
F4	Peak-to-RMS	$\frac{T_{PV}}{T_{RMS}}$	F8	Mean Frequency	$\frac{1}{N} \sum_{k=1}^N f_k$

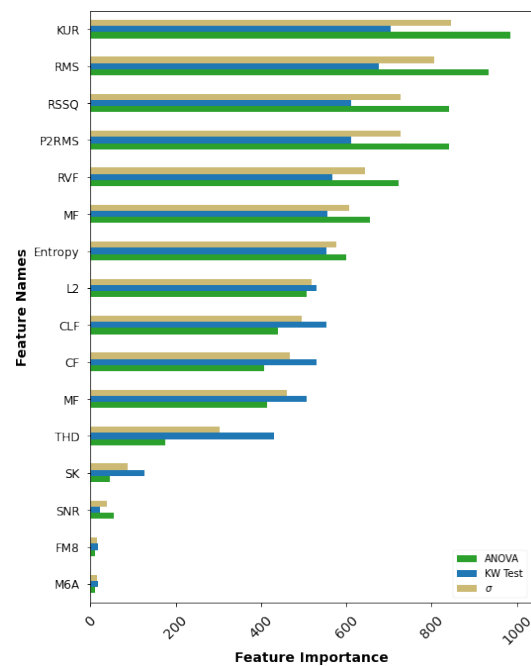


Figure 6. Feature selection based on σ score.

Figures 7–9 illustrate the vibration signals acquired at different health states of the gear-actuator system. Vibration data for a 1-s duration from all three health states (healthy, fault-1, and fault-2) are shown in the figures. Corresponding frequency domain representation is also shown using the fast Fourier transform (FFT) algorithm. Time and frequency representation gave us understanding of the signals acquired from different sensors. Sensor A acquired distinguishable signals in all three health states of the system. On the other hand, Sensor B and Sensor C captured fault frequencies; however, these were not as strong as those of the Sensor A signals.

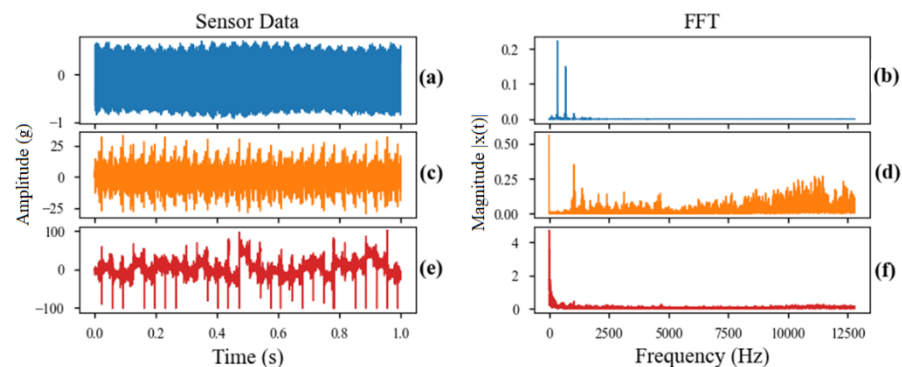


Figure 7. Sensor A signals. (a) Healthy, (b) Healthy FFT, (c) Fault-1 (d) Fault-1 FFT, (e) Fault-2, (f) Fault-2 FFT.

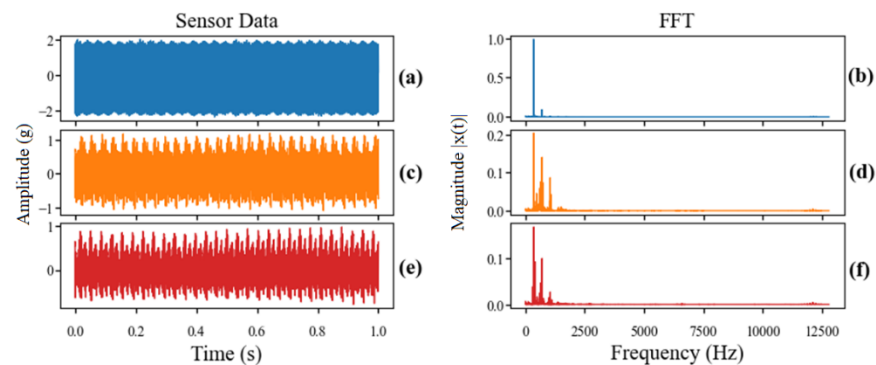


Figure 8. Sensor B signals. (a) Healthy, (b) Healthy FFT, (c) Fault-1 (d) Fault-1 FFT, (e) Fault-2, (f) Fault-2 FFT.

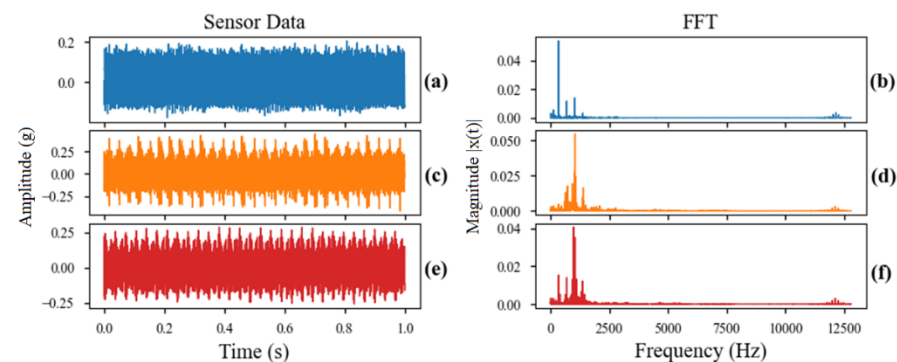


Figure 9. Sensor C signals. (a) Healthy, (b) Healthy FFT, (c) Fault-1 (d) Fault-1 FFT, (e) Fault-2, (f) Fault-2 FFT.

5.2. Failure Pattern Recognition

Selected features were further reduced to a lower dimension using four dimensionality reduction techniques. Feature space after reduction using these algorithms is presented in Figure 10. A fully connected ANN model was trained for all the dimension reduction algorithms, and the training metrics are presented in Figure 11. As the TSNE provided the highest accuracy and lowest loss score, this technique was adopted to further train the other sensor data. All the ANN model parameters were kept the same for all three sensors' acquired vibration data. The model was trained on a computer with AMD Ryzen 7 2700 octa-core CPU with 32 GB of RAM.

The deep learning architecture was supported by TensorFlow-GPU installed on NVIDIA GTX 970 with 4 GB VRAM. A list of model parameters is presented in Table 5. During training, 70% of the data are used for training and 30% are kept for the testing. To avoid overfitting in the model learning stage, a hyperparameter named "Dropout" was used for the hidden layers. Dropout randomly discards the effect of some neurons during the training stage to avoid local generalization. In our model, dropout rate was set to be 0.15. After the training was completed, the model was used to predict the test dataset. The confusion matrix for all three-sensor data is presented in Figure 12. It can be seen that Sensor A had the highest number of accurate predictions for all three classes. In the case of 4500 healthy, fault-1 and fault-2 data points, Sensor A predicted 4473, 3924, and 4391 data points accurately. On the other hand, Sensor B predicted 1667, 4015, and 3275 data points and Sensor C predicted 2309, 3714, and 2084 data points accurately for H, F1, and F2 classes, respectively.

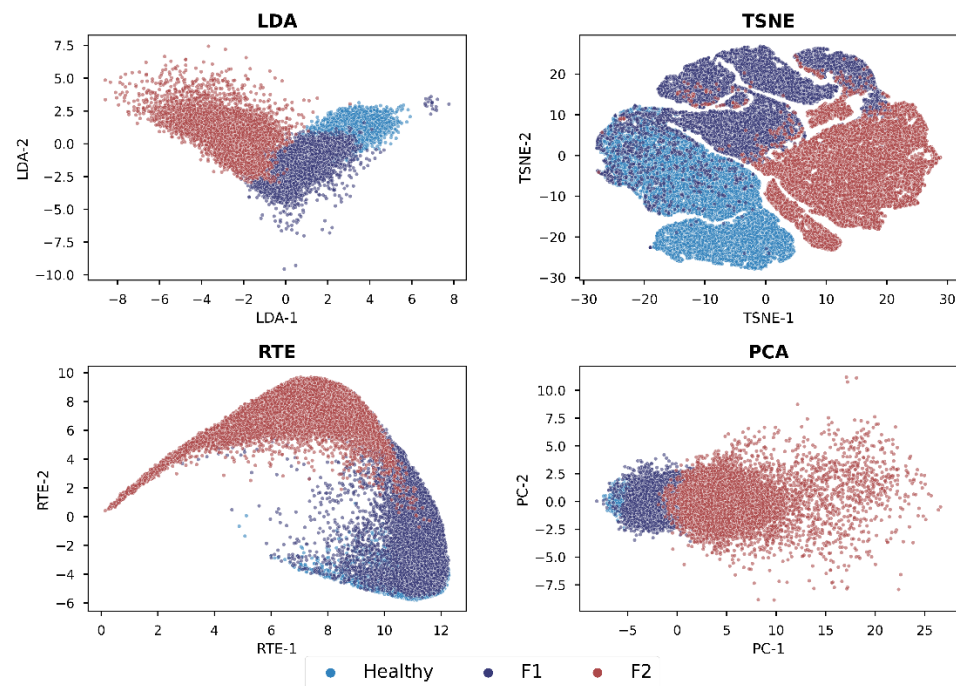


Figure 10. Reduced feature space using different algorithms.

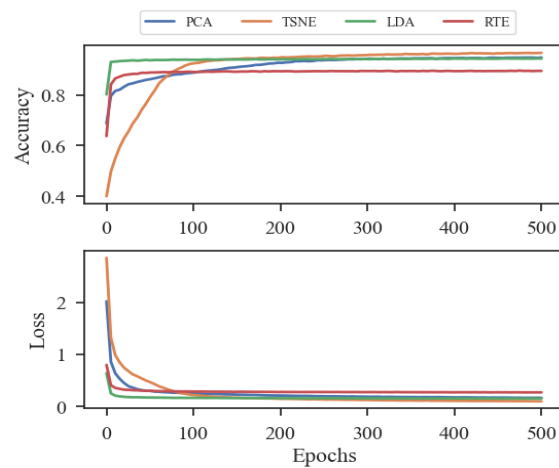


Figure 11. ANN training scores.

Table 5. Performance evaluation of different sensor data.

Sensor Position	Metrics	H	F1	F2
Sensor A	Precision	0.95	0.97	0.93
	Recall	0.99	0.87	0.98
	F1-Score	0.97	0.92	0.95
	Accuracy	0.95		
Sensor B	Precision	0.57	0.88	0.54
	Recall	0.37	0.89	0.73
	F1-Score	0.45	0.89	0.62
	Accuracy	0.66		
Sensor C	Precision	0.50	0.82	0.48
	Recall	0.51	0.83	0.46
	F1-Score	0.51	0.82	0.47
	Accuracy	0.60		

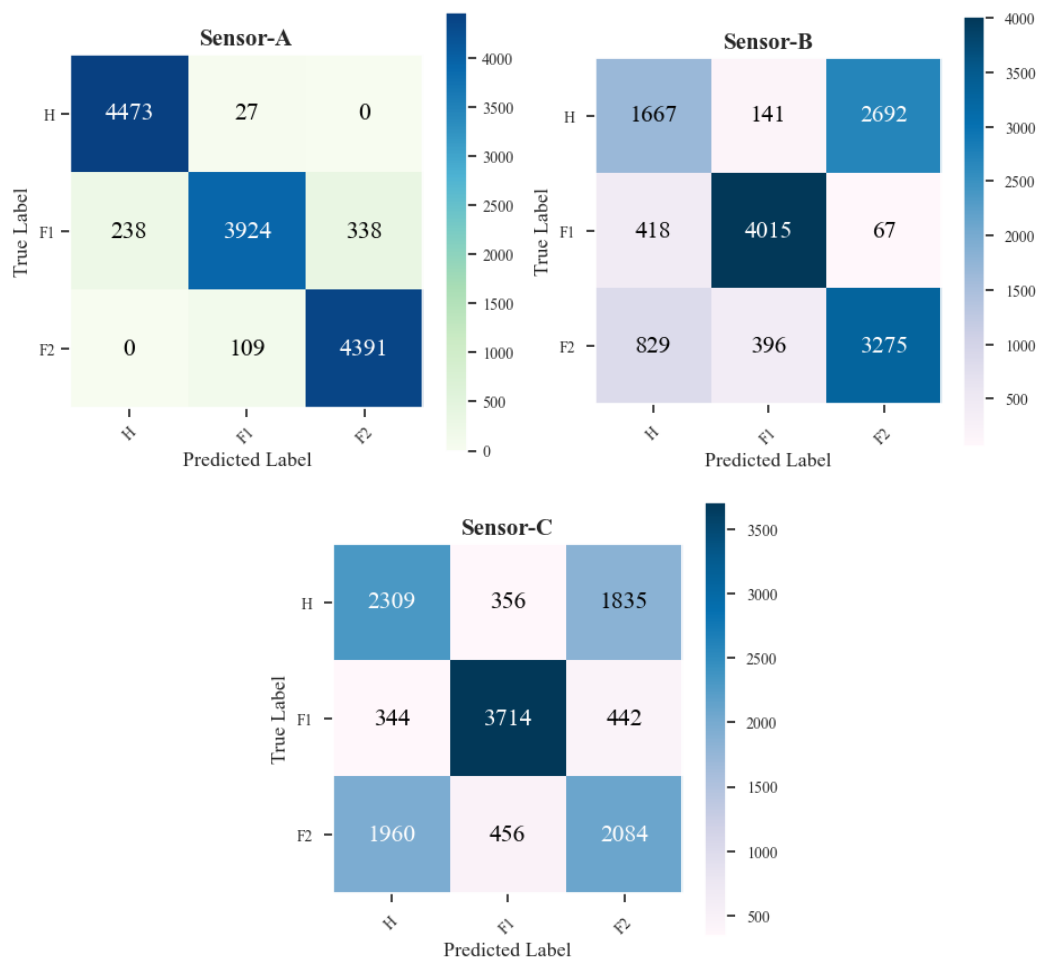


Figure 12. Confusion matrix for different sensors.

5.3. Validation of the Proposed Method

In a classification problem, analyzing only the accuracy score is not sufficient. To better understand the performance of the ANN model, we computed several performance metrics based on the predicted values and actual values. Mathematical expressions of the metrics are shown in (17)–(20):

$$\text{Precision, } P = \frac{TP}{TP + FP} \quad (17)$$

$$\text{Recall, } R = \frac{TP}{TP + FN} \quad (18)$$

$$\text{F1 Score, } F = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (19)$$

$$\text{Accuracy, } A = \frac{TP + TN}{TP + TN + FP + FN} \quad (20)$$

where,

TP = True positive

TN = True negative

FP = False positive

FN = False negative

Some “False” data can be classified as a “True” class in DL algorithms. For example, an F1 feature can be classified as H, and, hence, the total true predictions of H are increased. This does not necessarily mean that the model is performing well for the H class prediction.

To understand the underlying numbers of each class prediction, we compared all three-sensor data using the matrices computed (17)–(20). The comparison is shown in tabulated form in Table 5.

6. Conclusions

We propose a hybrid fault diagnosis approach of the gear-actuator system using a fusion of model-based and data-driven approaches. A FEM-based physical model is used to locate the appropriate sensor location through modal analysis. FFT and PSD data of initial random vibration were used as input to the FEM simulation and, based on the maximum mass deformation, the most suitable mounting point for the vibration sensor was obtained. Three different mounting positions were used for sensor positioning, including the best location found through FEM analysis. Several features from the vibration signals were computed and the best features were selected through a two-stage elimination-selection process. Later, an ANN model was used to learn the pattern of different health states of the gear-actuator system. Two more additional sensor locations were arbitrarily selected to compare the efficacy of FEM modeling. It was found that the FEM optimized sensor location was able to provide the maximum accuracy in the ANN FDI framework of the gear-actuator system. This study contributes the following:

- FEM modeling provides a robust analysis for sensor positioning through a detailed gear-actuator physical model.
- The ANN model is built on two different feature selection approaches ensuring the effectiveness of training data.

Finally, the combination of model-based and data-driven approaches provided a thorough understanding of the FDI of the gear-actuator system. At the same time, it preserved system dynamics for multi DOF operation, as well as ensured FDI in multiple operating conditions through deep learning modeling. This particular approach is advantageous for two reasons. Firstly, this approach does not require prior knowledge or a complicated mathematical model representation for the FEM analysis. Secondly, it considers multiple health indicators (features) in various domains. The feature selection technique can be updated, based on the system requirement and operating condition, which improves the prediction accuracy with lower computational cost.

In the future, we look forward to extending this model to predict future patterns by estimating the remaining useful life (RUL) of the gear-actuator system.

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

Funding: This research was supported by the MSIT (Ministry of Science and ICT), Korea, under the Grand Information Technology Research Center support program (IITP-2022-2020-0-01612) supervised by the IITP (Institute for Information & communications Technology Planning & Evaluation).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: We acknowledge LIG Nex1, Korea for their support.

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

References

1. Lei, Y.; Yang, B.; Jiang, X.; Jia, F.; Li, N.; Nandi, A.K. Applications of machine learning to machine fault diagnosis: A review and roadmap. *Mech. Syst. Signal Process.* **2020**, *138*, 106587. [\[CrossRef\]](#)
2. Liu, L.; Guo, Q.; Liu, D.; Peng, Y. Data-Driven Remaining Useful Life Prediction Considering Sensor Anomaly Detection and Data Recovery. *IEEE Access* **2019**, *7*, 58336–58345. [\[CrossRef\]](#)
3. Kim, N.-H.; An, D.; Choi, J.-H. *Prognostics and Health Management of Engineering Systems*; Springer International Publishing: Cham, Switzerland, 2017; pp. 18–138.
4. Wang, T.; Han, Q.; Chu, F.; Feng, Z. Vibration based condition monitoring and fault diagnosis of wind turbine planetary gearbox: A review. *Mech. Syst. Signal Process.* **2019**, *126*, 662–685. [\[CrossRef\]](#)
5. Krysanter, M.; Frisk, E. Sensor placement for fault diagnosis. *IEEE Trans. Syst. Man Cybern. Part A Syst. Hum.* **2008**, *38*, 1398–1410. [\[CrossRef\]](#)
6. Sarrate, R.; Nejari, F.; Rosich, A. Sensor placement for fault diagnosis performance maximization in Distribution Networks. In Proceedings of the 2012 20th Mediterranean Conference on Control & Automation (MED), Barcelona, Spain, 3–6 July 2012; pp. 110–115. [\[CrossRef\]](#)
7. Sztzyber, A. Sensor Placement for Fault Diagnosis Using Graph of a Process. *J. Phys. Conf. Ser.* **2017**, *783*, 012007. [\[CrossRef\]](#)
8. Wu, J.; Yang, Y.; Wang, P.; Wang, J.; Cheng, J. A novel method for gear crack fault diagnosis using improved analytical-FE and strain measurement. *Measurement* **2020**, *163*, 107936. [\[CrossRef\]](#)
9. Sapena-Bano, A.; Chinesta, F.; Pineda-Sanchez, M.; Aguado, J.; Borzacchiello, D.; Puche-Panadero, R. Induction machine model with finite element accuracy for condition monitoring running in real time using hardware in the loop system. *Int. J. Electr. Power Energy Syst.* **2019**, *111*, 315–324. [\[CrossRef\]](#)
10. Ezzat, A.A.; Tang, J.; Ding, Y. A model-based calibration approach for structural fault diagnosis using piezoelectric impedance measurements and a finite element model. *Struct. Health Monit.* **2020**, *19*, 1839–1855. [\[CrossRef\]](#)
11. Weili, L.; Ying, X.; Jiafeng, S.; Yingli, L. Finite-Element Analysis of Field Distribution and Characteristic Performance of Squirrel-Cage Induction Motor with Broken Bars. *IEEE Trans. Magn.* **2007**, *43*, 1537–1540. [\[CrossRef\]](#)
12. Vaseghi, B.; Takorabet, N.; Meibody-Tabar, F. Fault analysis and parameter identification of permanent-magnet motors by the finite-element method. *IEEE Trans. Magn.* **2009**, *45*, 3290–3295. [\[CrossRef\]](#)
13. Hoang, D.-T.; Kang, H.-J. A survey on Deep Learning based bearing fault diagnosis. *Neurocomputing* **2018**, *335*, 327–335. [\[CrossRef\]](#)
14. Fink, O.; Wang, Q.; Svensen, M.; Dersin, P.; Lee, W.J.; Ducoffe, M. Potential, challenges and future directions for deep learning in prognostics and health management applications. *Eng. Appl. Artif. Intell.* **2020**, *92*, 103678. [\[CrossRef\]](#)
15. 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. [\[CrossRef\]](#)
16. Alam Shifat, T.; Jang-Wook, H. Remaining Useful Life Estimation of BLDC Motor Considering Voltage Degradation and Attention-Based Neural Network. *IEEE Access* **2020**, *8*, 168414–168428. [\[CrossRef\]](#)
17. Alam Shifat, T.; Hur, J.-W. ANN Assisted Multi Sensor Information Fusion for BLDC Motor Fault Diagnosis. *IEEE Access* **2021**, *9*, 9429–9441. [\[CrossRef\]](#)
18. He, M.; He, D. Deep Learning Based Approach for Bearing Fault Diagnosis. *IEEE Trans. Ind. Appl.* **2017**, *53*, 3057–3065. [\[CrossRef\]](#)
19. Jing, L.; Zhao, M.; Li, P.; Xu, X. A convolutional neural network based feature learning and fault diagnosis method for the condition monitoring of gearbox. *Measurement* **2017**, *111*, 1–10. [\[CrossRef\]](#)
20. 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. [\[CrossRef\]](#)
21. Yan, X.; Liu, Y.; Jia, M.; Zhu, Y. A multi-stage hybrid fault diagnosis approach for rolling element bearing under various working conditions. *IEEE Access* **2019**, *7*, 138426–138441. [\[CrossRef\]](#)
22. Alam Shifat, T.; Hur, J.-W. EEMD assisted supervised learning for the fault diagnosis of BLDC motor using vibration signal. *J. Mech. Sci. Technol.* **2020**, *34*, 3981–3990. [\[CrossRef\]](#)
23. Judd, C.M.; McClelland, G.H.; Ryan, C.S. *Data Analysis: A Model Comparison Approach to Regression, ANOVA, and Beyond*; Routledge: London, UK, 2017.
24. Rao, S.S.; Atluri, S.N. The Finite Element Method in Engineering. *J. Appl. Mech.* **1983**, *50*, 914. [\[CrossRef\]](#)
25. Meyers, V.J.; Smith, I.M.; Griffiths, D.V. Programming the Finite Element Method. *Math. Comput.* **1989**, *53*, 763. [\[CrossRef\]](#)
26. Reddy, J.N. *Introduction to the Finite Element Method*; McGraw-Hill Education: New York, NY, USA, 2019.
27. Saravanan, N.; Ramachandran, K.I. Incipient gear box fault diagnosis using discrete wavelet transform (DWT) for feature extraction and classification using artificial neural network (ANN). *Expert Syst. Appl.* **2010**, *37*, 4168–4181. [\[CrossRef\]](#)
28. Samanta, B.; Al-Balushi, K.R. Artificial Neural Network Based Fault Diagnostics of Rolling Element Bearings Using Time-Domain Features. *Mech. Syst. Signal Process.* **2003**, *17*, 317–328. [\[CrossRef\]](#)
29. Reddy, G.T.; Reddy, M.P.K.; Lakshmana, K.; Kaluri, R.; Rajput, D.S.; Srivastava, G.; Baker, T. Analysis of Dimensionality Reduction Techniques on Big Data. *IEEE Access* **2020**, *8*, 54776–54788. [\[CrossRef\]](#)
30. Kalsoom, A.; Maqsood, M.; Ghazanfar, M.A.; Aadil, F.; Rho, S. A dimensionality reduction-based efficient software fault prediction using Fisher linear discriminant analysis (FLDA). *J. Supercomput.* **2018**, *74*, 4568–4602. [\[CrossRef\]](#)
31. Ng, A. Machine Learning Yearning. 2018. Available online: <http://www.mlyearning.org/> (accessed on 12 September 2021).
32. Bengio, Y.; Goodfellow, I.; Courville, A. *Deep Learning*; MIT Press: Cambridge, MA, USA, 2017; Volume 1.

-
33. Alam Shifat, T.; Hur, J. An Improved Stator Winding Short-circuit Fault Diagnosis using AdaBoost Algorithm. In Proceedings of the 2020 International Conference on Artificial Intelligence in Information and Communication (ICAIIIC), Fukuoka, Japan, 19–21 February 2020; pp. 382–387. [[CrossRef](#)]
 34. Yan, X.; Liu, Y.; Jia, M. A Feature Selection Framework-Based Multiscale Morphological Analysis Algorithm for Fault Diagnosis of Rolling Element Bearing. *IEEE Access* **2019**, *7*, 123436–123452. [[CrossRef](#)]