

Fault Diagnosis of Industrial Robot Bearings Based on Discrete Wavelet Transform and Artificial Neural Network

Alaa Abdulhady Jaber^{1,2} and Robert Bicker¹

¹*School of mechanical and Systems Engineering, Newcastle University, Newcastle upon Tyne, NE1 7RU, UK*

*a.jaber@ncl.ac.uk
robert.bicker@ncl.ac.uk*

²*Mechanical Engineering Department, University of Technology, Baghdad, Iraq*

ABSTRACT

Industrial robots have long been used in production systems in order to improve productivity, quality and safety in automated manufacturing processes. An unforeseen robot stoppage due to different reasons has the potential to cause an interruption in the entire production line, resulting in economic and production losses. The majority of the previous research on industrial robots health monitoring is focused on monitoring of a limited number of faults, such as backlash in gears, but does not diagnose the other gear and bearing faults. Thus, the main aim of this research is to develop an intelligent condition monitoring system to diagnose the most common faults that could be progressed in the bearings of industrial robot joints, such as inner/outer race bearing faults, using vibration signal analysis. For accurate fault diagnosis, time-frequency signal analysis based on the discrete wavelet transform (DWT) is adopted to extract the most salient features related to faults, and the artificial neural network (ANN) is used for faults classification. A data acquisition system based on National Instruments (NI) software and hardware was developed for robot vibration analysis and feature extraction. An experimental investigation was accomplished using the PUMA 560 robot. Firstly, vibration signals are captured from the robot when it is moving one joint cyclically. Then, by utilising the wavelet transform, signals are decomposed into multi-band frequency levels starting from higher to lower frequencies. For each of these levels the standard deviation feature is computed and used to design, train and test the proposed neural network. The developed system has showed high reliability in diagnosing several seeded faults in the robot.

Alaa Abdulhady Jaber et al. This is an open-access article distributed under the terms of the Creative Commons Attribution 3.0 United States License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

1. INTRODUCTION

The Robot Institute of America (RIA) has defined an industrial robot as a reprogrammable multifunctional manipulator designed to move material, parts, tools, or specialized devices through variable programmed motions for the performance of a variety of tasks (Spong et al., 2005). However, an unforeseen robot stoppage has the potential to cause an interruption in the entire production line, resulting in economic and production losses. Availability and maintainability, which can be defined as the probability of a system operating satisfactorily in any time period and its capability of being repaired, are therefore critical for industrial robots. Thus, the automated supervision of the robot system is desirable, as this can increase robot availability and maintainability and reduce operator effort. Industrial robots are extremely complex mechanism and hence the application of condition monitoring for them differs from that of 'simple' rotating machinery. This is basically due to the instantaneous change of geometrical configuration of the robot arm. Generally, there are two approaches to condition monitoring, which are model-based and model-free. Either of these approaches or a combination of both could be adopted in industrial robot condition monitoring. Filaretov et al. (1999) used a nonlinear model to address problems of fault detection and isolation in complex systems, such as in robot manipulators. Algebraic functions were implemented to design the nonlinear diagnostic observer, which was able to dispense with the linearization in nonlinear models to avoid model errors. The robot modeling was conducted using Matlab in discrete time. It was shown that, despite the fact that the use of this model dispenses with linearization, it does not allow some faults to be isolated. In terms of application model-based condition monitoring in gearboxes Recently, Liang et al. (2015) have developed a dynamic model to simulate the vibration source signals for a planetary gearbox in the healthy and the cracked tooth conditions. The signals were

analysed using time- and frequency-domain feature extraction techniques. For model verification experimental work was accomplished and the results were deemed acceptable. However, because precise mathematical models for complex systems like a robot are difficult to obtain, model-free methods based on artificial intelligence (AI) or statistical approaches have become prevalent choices for robot health monitoring.

Currently, there are few commercially available solutions that allow for the automated monitoring of the mechanical components of a robot, and therefore the ability to continuously monitor the status and condition of robots has become an important research topic in recent years and is now receiving considerable attention. The backlash in the power transmission system of a robot, which is defined as the clearance between the non-working flanks of the teeth of a gear pair when the working flanks are in contact (British-Standard, 2007), may cause torque variations. However, the electric motor itself generates what is known as a back electromotive force (EMF) when subjected to mechanical load making them acting as a torque transducer (Yuan et al., 2011). The torque variations measurements via current fluctuations on robotic actuators have been applied for robot monitoring (Abdul and Liu, 2008, Yuan et al., 2011). The advantage of this technique to the robots health monitoring is that the motor current can remotely be measured along the power cables utilizing standard current sensors without supplementary instrumentation on the robot. Some reported robot fault diagnostic systems are based on acoustic signals analysis. Such systems would have to be able to distinguish the correct information from the ambient noise. Case-based reasoning and signal processing were adopted to build an approach to diagnosis the faults in an industrial robot (Olsson et al., 2004). Wavelet analysis was applied to remove noise from the acoustic signals and to extract the most relevant features, which were then sent to the classification component, which uses case-based reasoning to identify the class of faults according to the characteristic of the previous fault cases. Experimental work on an industrial robot was used to assess the performance of this approach. The same principle was applied to an industrial robot, but on this occasion the artificial neural network (ANN) was used for noise analysis and classification (Yildirim and Eski, 2010). Noise sensors with data acquisition hardware and feature extraction software were used to prepare the training data for designing the ANN-based noise fault detection of robot manipulator's joints.

Vibration signal respond immediately to manifest itself if any change has appeared in the monitored machine (Anil Kumar et al., 2015). Vibration analysis approach is often used and represents the vast majority of utilized technique for industrial robot health monitoring, as it provides an easy and cost-effective sensing technique to detect faults in machines and for this reason it will also be used in this research. Pan *et al.* used vibration signals during normal operation to diagnose joint-backlash on a PUMA 762

industrial robot (Pan et al., 1998). Time-domain and frequency-domain analyses were employed to identify features such as probability and density. ANNs were then used for pattern recognition. One accelerometer was fixed to the robot end effector to measure vibration responses. Additionally, different levels of backlash were artificially contrived in the robot's joints to validate this method. Another research study used wavelet multi-resolution analysis (WMRA) coupled with a neural network-based approach in order to diagnosis faults in an industrial robot manipulator (Datta et al., 2007). A Matlab-Simulink environment was used to monitor the neural network classifier for a robot used in semi-conductor fabrication. It was concluded that the WMRA is excellent for data reduction and capturing the important properties of signals. On the other hand, two neural networks have been used to propose an algorithm for the online monitoring of two-link manipulators (Van et al., 2011). This approach focuses on identifying changes in robot dynamics due to faults. It was noted that this technique was able to provide estimates of fault characteristics. Recent work has been conducted to detect backlash in the PUMA 560 robot (Jaber and Bicker, 2014, Jaber and Bicker, 2015). A fault detection system using wavelet analysis was successfully designed based on National Instruments (NI) software and hardware. The wavelet transform was adopted for feature extraction and the extracted features showed high sensitivity to changes in joint backlash.

Generally, in the literature concerning industrial robot fault diagnosis it was established that the majority of them are focusing on gear backlash faults diagnosis, implying there are essential shortcomings in the diagnosis of other types of transmission faults, such as gear tooth wear and breakage and inner and outer race bearing faults. Thus, in this research it was decided to fill part of this gap by developing a condition monitoring system based on the combination of the discrete wavelet transform (DWT), for feature extraction, and ANN, for feature classification, to detect and diagnose the most common faults that could be progressed in the bearing of industrial robot joints. The detection and diagnosis of other robot joint's faults, such as gear teeth and backlash faults, on the same robot were presented in (Jaber and Bicker, 2016, Abdulhady Jaber and Bicker, 2016). However, the main goal behind this work is to build a condition monitoring system that can be able to detect and diagnosis different mechanical and electrical faults in the robot.

2. DISCRETE WAVELET TRANSFORM (DWT)

Robot arms exhibit very complex dynamic behavior, and different defects can affect this behavior. Also, their motion is completely different from that of rotating machines (or other continuously moving machines), for which the majority of present CM systems have been designed. However, in order for the robot to complete any assigned

scenario, each joint in the robot body will move at different angular speeds (and accelerations), needing different torques, and rotating at different angles. When this motion is compared with that of rotating machines, it can be noticed that the latter move at continuous and fixed speed, while the former's movement is discrete and varying over time. Continuously rotating machines emit continuous signals during their operation. Therefore, the CM system will be relatively easy to implement as long as there is a continuous signal during any time period. On the other hand, the signals emitted from the robot's mechanical or electrical parts will be transitory and last for a very short time. The signals from faulty parts have a non-stationary nature. However, if the frequency component of the non-stationary signals is calculated using the Fourier transform, the results will represent the frequency composition averaged over the duration of the signal (Sawicki et al., 2009). Consequently, the characteristics of the transient signal cannot be described adequately using the Fourier transform, however, time-frequency analysis has been investigated and applied for the fault diagnosis of machinery because of its capability of signal representation in both the frequency and time domains (Sawicki et al., 2009, Al-Badour et al., 2011). This unique feature of time-frequency analysis techniques means that it is suitable for non-stationary signals. Moreover, time-frequency methods can give interesting information with regard to energy distribution over frequency bands. A number of techniques of time-frequency analysis, such as the short time Fourier transform and discrete wavelet transforms (DWT), have previously been used for fault detection and diagnosis. DWT is superior to short time Fourier transform, as the former has varying window lengths, and represents the signal as a sum of wavelets at different scales; therefore, it was used here.

The concept of the DWT is that filters with different cut-off frequencies are utilized to analyze the signal at different scales (Boukabache et al., 2013). Firstly, the signal is passed through a high-pass filter to analyze high frequencies, and then it is passed through a low-pass filter to analyze low frequencies. Generally, by using the DWT, a multi-resolution analysis can be performed at different frequency bands with different resolutions by decomposing the time domain signal (Sawicki et al., 2009, Debdas et al., 2011). Two sets of functions are employed in the DWT. These are called the wavelet function and the scaling function, which are associated with the high-pass (HP) and low-pass (LP) filters respectively. At the first level, the original signal $x[n]$ is decomposed by passing it through both of these filters and emerges as two signals, each one having the same number of samples as the original signal, and are termed as coefficients. In order to keep the total number of coefficients in the produced filtered signals equal to the original signal samples they are then down-sampled by a factor of 2, by keeping only one sample out of two successive samples. Thus, the extracted signal coefficients from the HP filter and after down sampling are called the

detail coefficients of the first level (cd_1). These coefficients contain the high frequency information of the original signal, whilst, the coefficients that are extracted from the LP filter and after the down sampling process are called the approximation coefficients of the first level (ca_1). The low frequency information of the signal is hidden in these coefficients. This can be expressed mathematically as (Vivas et al., 2013):

$$y_{high}[k] = \sum_n x[n] * g[2k - n] \quad (1)$$

$$y_{low}[k] = \sum_n x[n] * h[2k - n] \quad (2)$$

where $y_{high}[k]$ and $y_{low}[k]$ are the outputs of the high-pass and low-pass filters respectively, after down-sampling by 2. After obtaining the first level of decomposition, the above procedure can be repeated again to decompose ca_1 into another approximation and detail coefficients. This procedure can be continued successively until a pre-defined certain level up to which the decomposition is required to be found. At each decomposition level, the corresponding detail and approximation coefficients have specific frequency bandwidths given by $[0 - F_s/2^{l+1}]$ for the approximation coefficients (ca_l) and $[F_s/2^{l+1} - F_s/2^l]$ for the detailed one (cd_l) where F_s is the sampling frequency (Sawicki et al., 2009, Vivas et al., 2013). However, at every level, the filtering and down-sampling will result in half the number of samples (half the time resolution) and half the frequency band (double the frequency resolution). Also, due to the consecutive down sampling by 2, the total number of samples in the analyzed signal must be a power of 2 (Ghods and Lee, 2014). By concatenating all coefficients starting from the last level of decomposition, the DWT of the original signal is then produced, and it will have the same number of samples as the original signal. A schematic diagram illustrates how the multi-level decomposition is performed shown in Figure 1. The number of decomposition levels is identified by the lowest frequency band needed to be traced, and a higher number of decomposition levels are required if very low frequency band is investigated. However, the highest decomposition level that can be achieved is up to that the individual details consist of a single sample (Misiti et al., 1997).

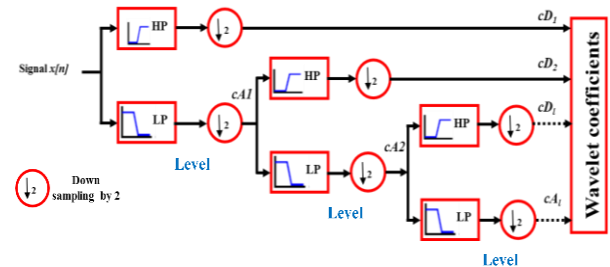


Figure 1. Signal decomposition using discrete wavelet transform

Once the approximation and detail coefficients are computed to different levels of decomposition, it becomes possible to reconstruct the approximation and detail signals at each level, in order to extract features, such as standard deviation, related to the frequency bands in each level. Each signal, however, will have the same number of samples as the original signal but with a definite frequency band. This can be achieved by up-sampling the approximation (or details) coefficients by two, since they were produced previously by down sampling by 2, and then passing them through high- and low-pass synthesis filters. In this work and based on a preliminary robot vibration analysis was established that eight decomposition levels are sufficient for analysing the robot vibration signals, as explained later.

3. ARTIFICIAL NEURAL NETWORK (ANN)

The ANN is computational structure inspired by the data processing and learning ability of biological neurons in the nervous system. It has a summation type of structure that consists of several layers of artificial neurons, which integrate the functionality of both memory and computation, designed to emulate the biological neurons. Each input signal flows through a gain or weight, called a synaptic weight or connection strength and can be positive or negative, integer or non-integer. Figure 2 illustrates a typical representation of an artificial neuron, with connection weights.

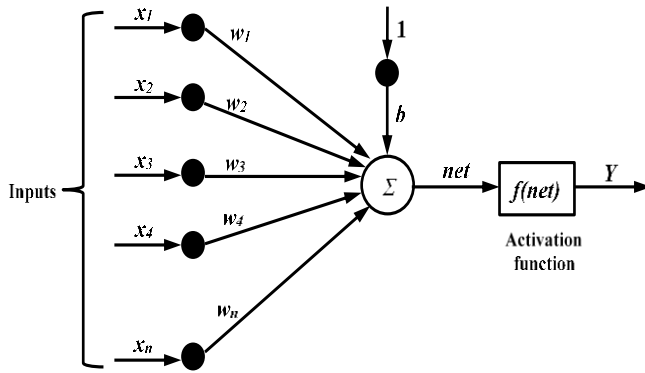


Figure 2. Artificial neuron with connection weights

Where $x_1, x_2, x_3, \dots, x_n$ represent the input vector to the neuron with associated weights $w_1, w_2, w_3, \dots, w_n$. The bias, b , (sometimes called threshold) is often connected to the neuron and it introduces an offset to the transfer function so that even if the input were zero the neurons would still have an output (Lim., 2009). Typically, the bias is set to a value of 1. However, the input to the neuron, which is called net , will be the summation of the multiplied inputs with their corresponding weights, and can be written as:

$$net = \sum_{i=1}^n w_i x_i = w_1 x_1 + w_2 x_2 + \dots + w_n x_n + b \quad (3)$$

where n is the number of inputs to the neuron. The resultant net is then passed through an activation function ($f(net)$), sometimes called the transfer function that can be linear or nonlinear, to produce the output which can be represented as:

$$Y = f(net) = f\left(\sum_{i=1}^n w_i x_i\right) \quad (4)$$

Neural networks come in many forms and the structure of the connections between neurons and the computation it performs vary between the different neural models. ANNs are generally classified as feed-forward, e.g. the multilayer perceptron neural network (MLPN), and feed-back (or recurrent) types, such as Hopfield network. In the feed-forward class, the signals travel only in the direction from the input to the output, whereas in recurrent neural network (RNN), the signals can flow in the forward as well as backward or lateral direction. However, the main characteristic of a neural network is that, when the training process is performed well, the mapping formed by the network can show its capability for generalization beyond the training data and not to memorise the training data. The majority of fault diagnosis applications utilizing ANNs utilize a feed-forward architecture, with the MLPN being most popular and thus it adopted for use in this research.

4. THE EXPERIMENTAL SET-UP

The experimental work has been performed using the PUMA 560 robot, which is a PC-controlled serial manipulator with six revolute joints/degrees of freedom (DOF). It is an old (>25 years), but functional, multi-joint robot and was used in this project, due to the availability of the spare parts that are required for simulation of the different faults. Three single axis, analogue outputs, 70g, MEMS accelerometers type ADXL001 from Analog Device have been used for signal capture. These accelerometers have been put together in a configuration that allows the vibration signals in the X, Y, and Z directions to be measured using a specially designed adapter, as shown in Figure 3.

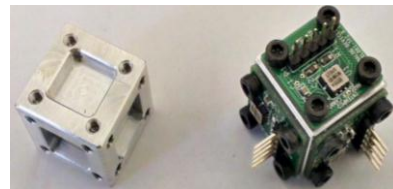


Figure 3. Designed 3 axis accelerometer adapter

It may be assumed that, to detect the abnormalities in a machine, the sensors should be located near expected damage locations. However, there are cases where the damage is more recognizable at other locations on the machine. So, prior to acquiring the data for fault detection, it is crucial to locate the sensitive positions on the robot. A preliminary analysis was performed and it was found that the best position to fix the accelerometer on the robot is near to the wrist joint, as shown in Figure 4. The signals from the accelerometers are fed to a 14-bit NI data acquisition card (DAQ) type USB-6009. However, the utilized accelerometers were calibrated against a conventional piezoelectric accelerometer before they are being used for robot monitoring.

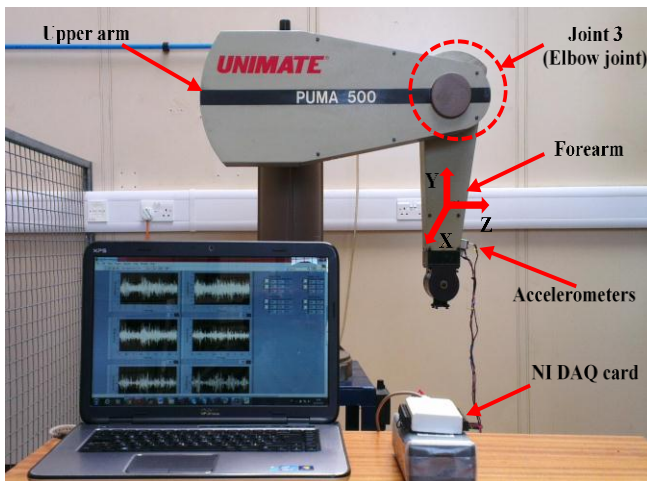


Figure 4. The experimental set-up

In terms of data acquisition software, many functions have been provided by LabVIEW, such as the interface with Matlab and C language, which make graphical programming in LabVIEW more flexible. Furthermore, Matlab software has many specialized, efficient, and easy-to-learn toolboxes, and so some of comments in the wavelet analysis toolbox are used here. This toolbox offers many options in helping to accomplish wavelet analysis, such as specifying the number of decomposition levels and selecting a suitable wavelet family. Therefore, combining the capabilities of Matlab in data processing and the advantages of LabVIEW graphical programming, the virtual instrument produced will have more capabilities and greater flexibility. In order to achieve the LabVIEW and Matlab interface, LabVIEW offers a Matlab script node to call a programme written in the Matlab language, which was used in this work. A sample from the prepared LabVIEW code is shown in the Figure 5. The LabVIEW block diagram begins with a *for loop*. However, many parameters in the code have to be tuned properly for accurate wavelet analysis, such as which wavelet family, number of decomposition levels, and this was achieved using (**wavedec**). Researchers who want to accomplish feature extraction using wavelet are advised to

read in detail the (wavelet toolbox user guide) (Misiti et al., 2001).

In the experimental process and since the faults are simulated in joint 3, as discussed in the following section, the robot was programmed to oscillate over its full range of movement (270 degree), as some of the robot faults may exhibit direction-dependent symptoms (Datta et al., 2007). This is also in order to relate the produced vibration to the moving joint's transmission components such as gears and bearings. Vibration signals from the three axes of measurements are captured at a sampling frequency of 1031 Hz with sample size equal to 4096. This frequency is about ten times higher than the 11th natural frequency of the robot, based on a conducted preliminary robot modal (frequency response function) analysis, which fulfils the sampling theory that requests the sampling frequency to be at least two times higher than the highest system frequency (Mohanty, 2015). Thus, eight wavelet decomposition levels are found to be sufficient for analysing the robot vibration signal. Statistical features are computed from each detail signal produced after the DWT analysis; these features are utilized for designing a neural network for fault diagnosis.

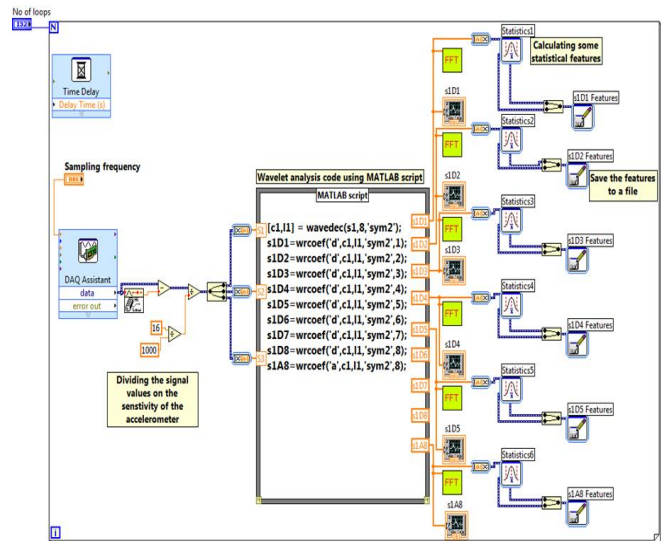


Figure 5. Part from the prepared LabVIEW code

5. ROBOT'S MECHANICAL CONSTRUCTION AND FAULTS

Joints in industrial robots are commonly actuated by electrical motors. Permanent magnet servomotors are a popular choice to produce the driving force to move robot joints because of their easy operation, high power density and performance (Halme, 2006). In general, servomotors are electromechanical components in which faults can originate for electrical, mechanical and other external reasons. In order to transform motor power to the robot joints, mechanical reduction gears in the transmission system are normally used. The power is then transmitted from the input to the output shaft through the gear contacts and mesh.

However, this paper concentrates on fault detection in the elbow joint (joint 3) of the robot, which is shown in Figure 3 above. The mechanical construction of this joint is illustrated in Figure 6; it consists of a two-stage gear train system housed in the end of the upper arm and connected to a DC motor by a drive shaft. There is a bevel pinion on the input shaft meshed with a bevel wheel on one end of an idler shaft. Also, a spur pinion at the other end of the idler shaft engages with a spur wheel fixed to the forearm, and therefore rotates the forearm around the elbow axis. A number of deep groove ball bearings are used to carry the input and idler shafts.

The main fault mechanisms that may appear in the joint gearbox are basically the same as those arising in other types of gearboxes, such as gear and bearing faults. The most common types of gear tooth failure are scuffing, cracking, macro- and micro-pitting, wear, bending fatigue, and fracture due to overload as well as backlash between mating teeth. Moreover, since gears are normally supported on rolling element bearings, faults in these bearings such as wear in the inner or outer races represent another typical type of fault in the robot's transmission system. However, the focus of this paper is limited for diagnosing only the bearing faults.

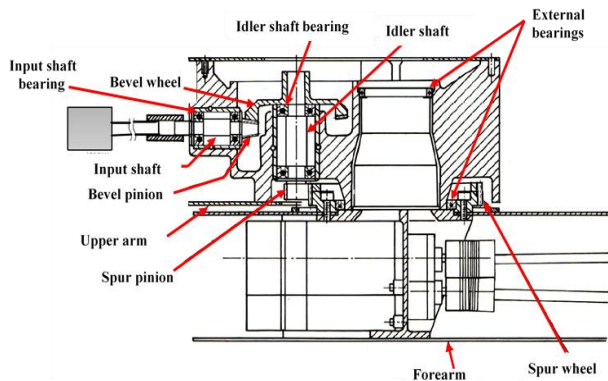


Figure 6. Schematic diagram of joint 3 (elbow joint)

5.1. Bearing Fault Simulation

It is accepted that the best way to introduce a fault in the robot is by operating the robot to execute a specific task for a considerable amount of time until a fault is developed. This method, however, has not been followed in this work primarily because it may result in other joints of the robot failing. The faults that will be introduced here are not truly representative of real faults in the robot, since in a practical situation the majority of faults in gears and bearings are progressive. However, this methodology has important advantages such as different types of fault can be produced in different components of the joint 3 gearbox and the degree of fault severity can be controlled.

In Figure 6 above, it can be observed that the input shaft is supported on two bearings of the same type. The one located

on the right, adjacent to the bevel pinion, is used for the bearing faults simulation. From the figure it may also be concluded that the applied load on these two bearings is not equally distributed; so, for future studies researchers can investigate the influence of the faulty bearing location on fault detectability. Fortunately, after dismantling the robot joint it was found that identical bearings are still available and a number of NSK bearings type N609 were purchased and utilized for faults seeding. Two fault types are created in the bearing, inner and outer races faults, using an electrical discharge machining (EDM) technique. However, due to the small size of the bearing, it was not possible to dismantle the bearing, in order to introduce a ball fault, without damaging it, which affects the fault size controllability of the inner race. The seeded inner race fault has 1 mm width, extending along the bearing width and as deep as the thickness of the inner race. Faults of two degrees of severity are introduced on the outer race of the bearing. Both of them are circular in shape, extended along the outer race thickness and have 1 mm and 2 mm diameters, respectively. The purpose of this variation in the outer race fault size is to test the proposed system in distinguishing the severity of different bearing faults. Figure 7 shows the healthy and the three faulty bearings.

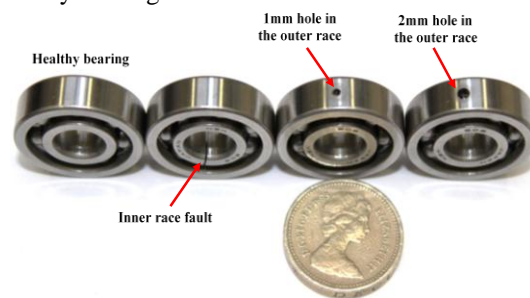


Figure 7. Healthy bearing and three faulty bearings with different fault types

6. ROBOT VIBRATION ANALYSIS

Selected results of the multi-resolution signal analysis using DWT extracted from the LabVIEW code are shown in Figures 8. This figure represents the analyzed signal from the X-axis accelerometer when the severest fault category, a 2 mm hole in the bearing outer race, is introduced in the robot. There are 10 sub-figures for each measurement axis, as in Figure 8; the first sub-figure is the original signal before the DWT is applied, while the rest are the eight constructed sub-signals from the detail coefficients (D1, D2, D3, D4, D5, D6, D7, and D8), and one from the approximation coefficient (A8), where each one contains a specific frequency band presented in the upper left corner of each sub-figure. These bands were computed based on the formulas that were presented in Section 2. Generally, accelerometers provide mixed information about vibration and movement acceleration in the robot (Rodriguez-Donate et al., 2010). The vibration signal is made up of high

frequencies, whereas movement acceleration signals are associated with low frequencies (Rodriguez-Donate et al., 2010). After conducting the DWT on the original signals and by investigating the produced sub-signals it was concluded that the approximation signals (A8) in the three axes are mostly related to the robot movement and not to its vibration components. Consequently, the approximation signals do not carry any useful information regarding robot faults and are not considered important in the present work. However, extracting the robot movement acceleration using DWT could be employed in future work to estimate the kinematics of the industrial robots (Rodriguez-Donate et al., 2010).

The sub-signals from the DWT analysis could be directly used as inputs to the classifier, which will be the ANN. However, feeding huge amounts of data, represented by 4096 samples from each sub-signal, into the classifier will make the classification process more complex and computationally intensive. To overcome this, a post processing stage is needed to take out significant parameters from the sub-signals, called the feature extraction stage, and will significantly reduce the data set size required to be passed to the classifier for correct fault diagnosis. Based on this, several statistical functions, e.g. the root mean square (RMS), standard deviation (STD), and kurtosis, are applied to the detail signals from the three axes of measurement, since approximation signals are neglected, as discussed earlier. These features are extracted in real-time after the data is captured and DWT analysis is applied using the designed LabVIEW code. However, if all computed features are used for fault classification, this will also involve significant amount of data for the classifier. Thus, only the most sensitive feature was utilized. The STD feature, which is normally used as a measure of extent of variation of the processed data and has the same units as the data, was found to be the most faults sensitive; and hence it was selected for comparison among fault severities. Therefore, 8 STD features from each axis signal are obtained, and by joining all features from the three axis signals together the input vector to the classifier will be constituted from 24 features.

Figure 9 shows the computed STDs from the eight detail signals from accelerometer axes X, Y, and Z when the three bearing faults are simulated. In this figure can be seen that STD vectors in Y-axis are squeezed except at higher frequency bands (D1 to D3) they are a bit diverged. However, the Y-axis results will not be ignored and will be included in the input vectors to the classifier, as sure the STDs in Y-axis will give a clear indication about a fault if it is developed, for example, in other joints or parts. The STD values in X- and Z-axis are increased as the fault severity increased and they are well separated except at the lower frequency bands from D6 to D8, this could be attributed to that their frequency ranges are not affect by the simulated faults. Generally speaking, if the STD vectors are very well isolated, the problem of fault classification will be easier

when ANN is used, the false diagnosis rate is reduced, and also the designed ANN will be of small size helping in lowering the required processing power.

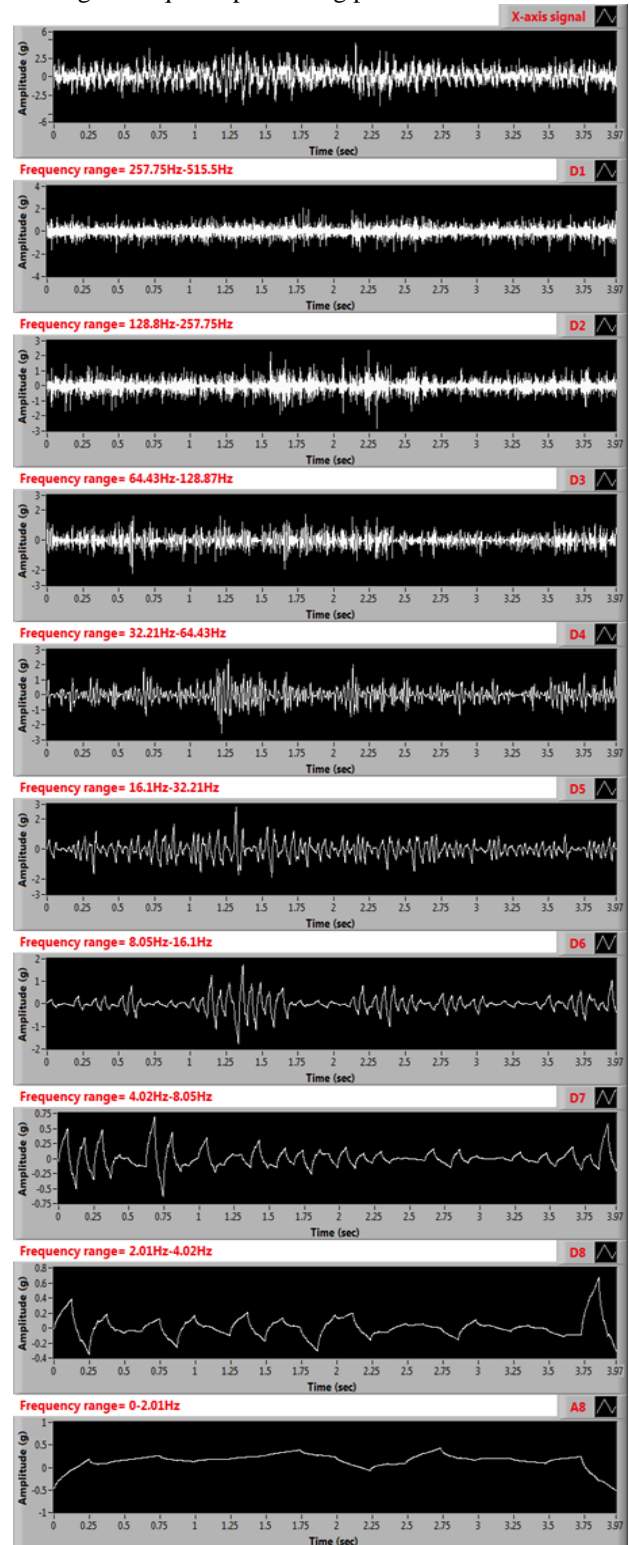


Figure 8. DWT analysis of the X-axis vibration signal from the robot when outer race 2mm hole fault simulated

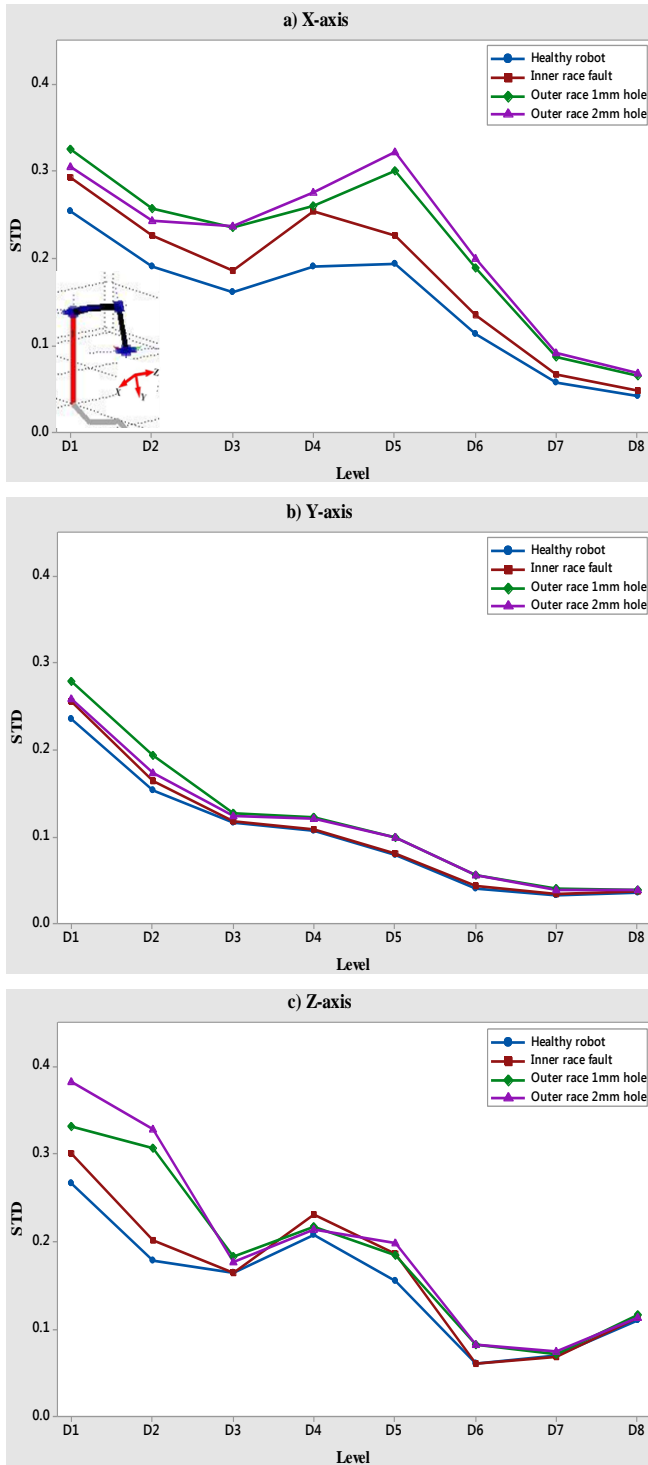


Figure 9. Standard deviation (STD) in all 8 detail signals when bearing faults simulated

7. DESIGN OF MULTILAYER PERCEPTRON NEURAL NETWORK (MLPN) FOR FAULT DIAGNOSIS

To model the nonlinear characteristics of the robot fault conditions, as in the work here, there are some issues

regarding the number of hidden layers, number of neurons in each layer, and also the activation functions in the hidden and output layers need to be considered. However, the number of input neurons depends on the number of measurements or the number of features extracted from each sensor signal. In this work, as mentioned early, the input features to the network are the standard deviation values of the detail signals of the wavelet analysis (D1 to D8), since the last level of approximation coefficient (A8) was found does not give useful information regarding robot fault. Therefore, the number of features will be 8 from each sensor (axis) signal with a total of 24 values (from the three axes) input to the network. The number of the robot health conditions to be classified identifies the required number of neurons in the output layer, which are four conditions in this work. Consequently, the number of MLPN input and output neurons are fixed (24 and 4 respectively). For any nonlinear system identification problem, at least one hidden layer is required (Negnevitsky, 2005). Additionally, a nonlinear, differentiable activation function for the hidden layer is needed (Mazumdar, 2006). For this purpose, the sigmoidal activation function has been utilized, since it is suitable for applications whose desired output is between 0 and 1, which is the case in this research (Pandya et al., 2012). Also, a linear function, which is normally used in the input and output layers, was used in the output layer. The only variable then remaining is the number of neurons in the hidden layer.

The appropriate selection of the number of hidden neurons is based on a balance between output accuracy and network size. Generally, there is no explicit mathematical or theoretical foundation to determine the best number of neurons in each hidden layer without training several networks and estimating the error of each network. There are several books and articles that offer a “rule of thumb” to select the optimal number of neurons in the hidden layer to be a starting point (Heaton, 2008) (S.N.Sivanandam. et al., 2006). Another possibility is by starting with a small number of neurons and gradually increasing them until little or no improvement is observed in the network performance. Therefore, after many trials it was found the optimum number of neurons that gives a compromise solution for the robot fault diagnosis problem is 17.

Matlab's neural network toolbox was utilized to design and test the network, based on the supervised, back-propagation learning technique. Many functions are provided by the toolbox for designing, training, visualizing and simulating the proposed neural networks, and numerous built-in functions for common neural network applications such as pattern recognition and clustering are included. Furthermore, the code for these functions can easily be modified to suit specific requirements. The block diagram in Figure 10 shows the main design steps for the ANN. Some of these steps are adjusted manually before running the prepared Matlab code, such as identifying the network

structure, activation functions and number of epoch, whereas others are performed automatically based on the code sequence, i.e. data loading and normalization, and weights and biases initialization.

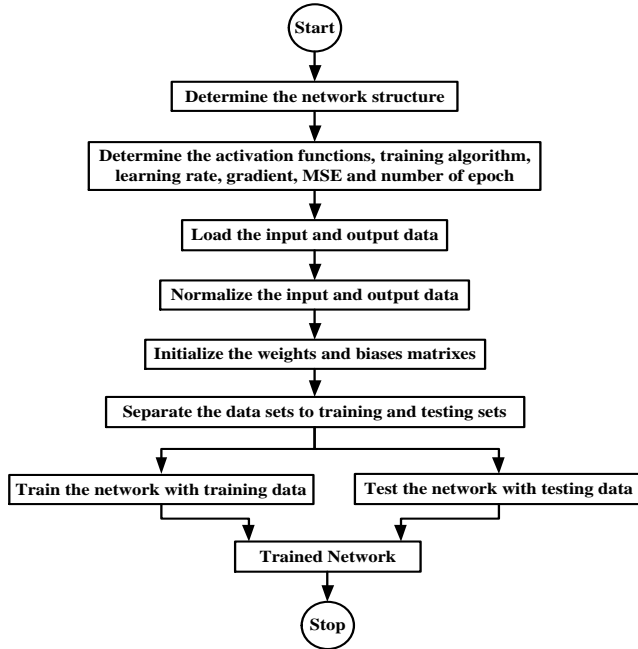


Figure 10. Block diagram of ANN design steps

The neural network input feature vectors consisting of one group representing the healthy state as well as three groups for different fault types. These groups are related to the bearing faults: inner race and 1 mm and 2 mm holes in the outer race. The data sets sizes of 100 samples for each health condition are used to train the network. After setting the network structure and loading the data, the weights and biases of the network are initialized depending on Matlab's random number generation function. The function *dividerand* was used to divide the data into three subsets; training, validation and testing sets with ratios of 70%, 15% and 15%, respectively. The parameters used for the network are shown in Table 1. The used data for training the ANN normally comes in a different range; therefore normalisation of the data is needed to avoid high values from being too dominant and to suppress the influence of the smaller ones (Subbaraj and Kannapiran, 2014, Pandya et al., 2012), this also will prevent weights from becoming too large which can lead numerical overflow. So, if all of the input data is normalized to be between 0 and 1 or -1 and 1, then the ANN will give equal priority to all inputs. In this work the data are normalized between 0 and 1 using the following equation:

$$x_n = \frac{x - x_{min}}{x_{max} - x_{min}} \quad (5)$$

where, x_n is the normalized value, and x_{min} and x_{max} are the minimum and maximum values among all values of the data. Additionally, the training can stop according to any one of the criteria of mean square error (MSE), the number of epochs of training or minimum performance gradient reaches certain value set as shown in Table 1. The MSE is calculated using Equation (6) below, where l is the total number of output neurons, t_k is the target output and y_k is the network output. After the training is completed and new input data sets have to be evaluated, which have to be normalised too as the network has been trained in this way. Also, the output results from the network need to be denormalised using the inverse of the normalization function.

$$MSE = \frac{1}{l} \sum_{k=1}^l (t_k - y_k)^2 \quad (6)$$

Number of input layer neurons	24
Number of hidden layer neurons	17
Number of output layer neurons	11
Number of hidden layers	1
Hidden layer activation function	Sigmoid
Output layer activation function	Linear
Training algorithm	To be identified
Learning rate	0.05
MSE stopping criteria	10e-4
Minimum performance gradient	10e-5
Maximum number of epoch	50000

Table 1. The used parameters for designing the neural network

The ANN learning process and the performance plot is depicted in Figure 11, which shows that the training with 487 epochs met the MSE stopping criteria (MSE less than 10E-4). Furthermore, it was established that the correlation coefficient (R) between the actual and desired (target) outputs has a value above 0.99 for the training, testing and validation data sets, which is indicative of a strong relationship between the outputs and targets of ANN. Thus, this network will be implemented for robot fault diagnosis.

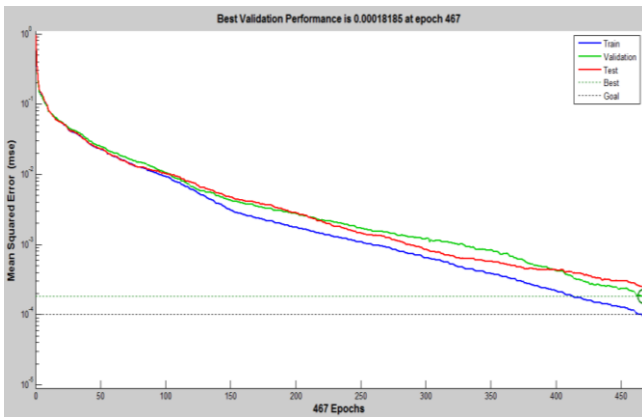
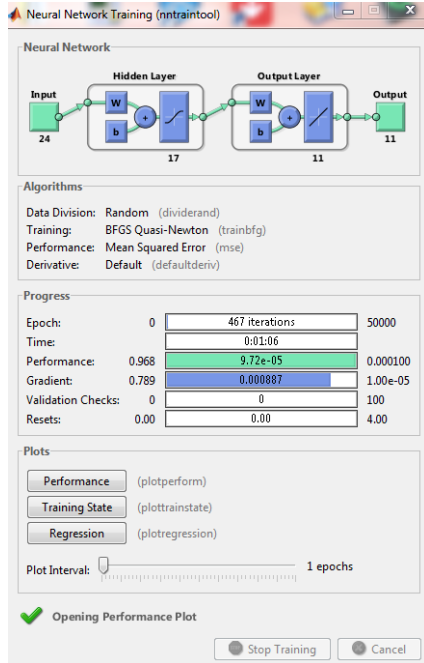


Figure 11. The ANN training process and performance plot

8. PERFORMANCE TESTING OF THE DESIGNED SYSTEM

With the ANN's design satisfying the requirements and the training completed, its performance needs to be tested using sets of data not previously presented to the network, and however, the well-trained network must be able to show its capability in classifying the unseen data samples. The ANN has been trained in a way that produced results from the first neuron in the output layer of the ANN representing the healthy operating condition of the robot, while the remaining three neurons are mapped to represent the three simulated faults on the robot, and thus a value of 1 for each output neuron is considered the target value. For instance, for the healthy case, the only output neuron that would have a value of 1 is the first one; therefore, it will be [1, 0, 0, 0]. The second neuron would take a value of 1 if the first fault type (inner race fault) is presented and thus the outputs will be [0, 1, 0, 0], and so on with the other fault types, as

illustrated in Figure 12. To test the network, a total of 80 unseen data samples were used. The first 20 samples represent the healthy condition of the robot while the remainder of the samples are divided into three groups of 20 samples each, corresponding to simulated robot faults. It is worthy to mention here that due to difficulty of dismantling and assembling the robot again each fault type was simulated in the robot one time only, but many data sets were captured, to compensate this limitation.

The test results are presented in Figure 13. The tests showed that the network efficiency in term of its capability in classifying the five different types of the robot health situation is 100 per cent, and the ANN design can differentiate the different faults of the robot with very good accuracy when confronted with unseen data. The proposed condition monitoring approach was developed to monitor a rather old, but operational, PUMA 560 robot, and thus it follows that more experimental work is needed to further evaluate its reliability in fault detection and diagnosis with the robot executing a variety of different tasks. Also, it should be tested on another PUMA 560 robot in order to verify that it can confidently be utilized with this model of robots. Of course it would be much more appropriate if the designed system is generalized for any industrial robot model; however, a considerable effort and time are needed to investigate the required scientific analysis methods, and the appropriate approaches for hardware and software development.

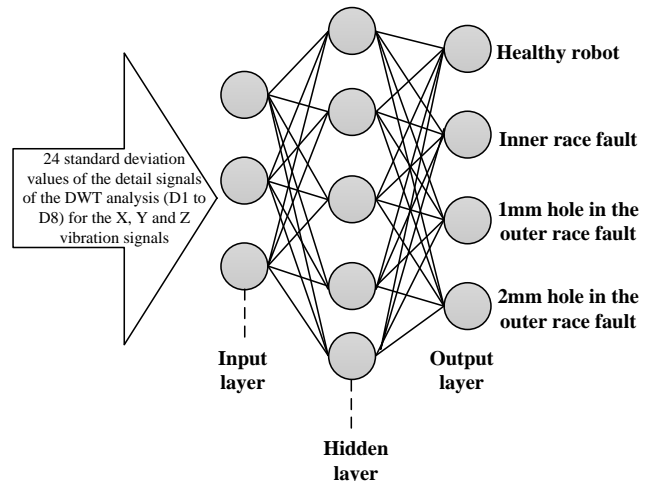
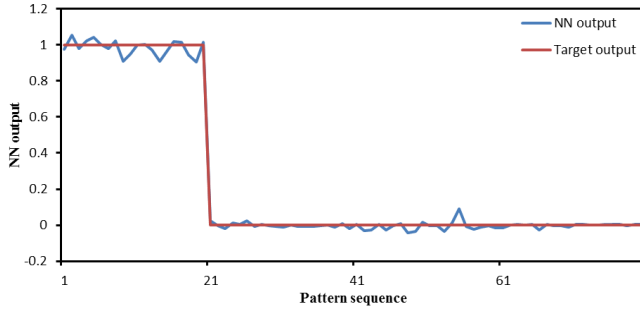
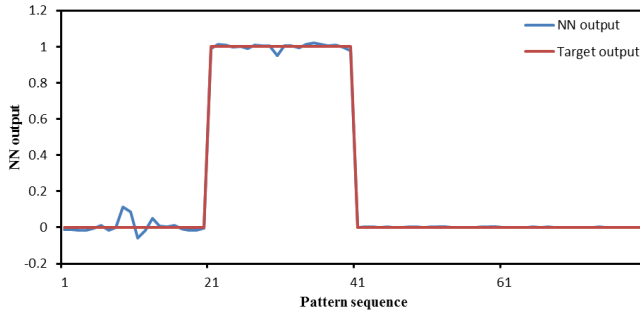


Figure 12. Schematic diagram of the designed MLPN

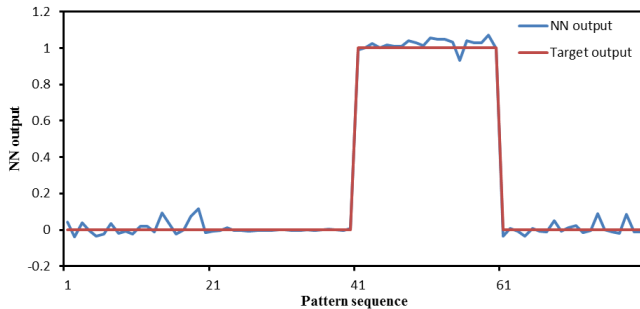


a) Robot healthy condition

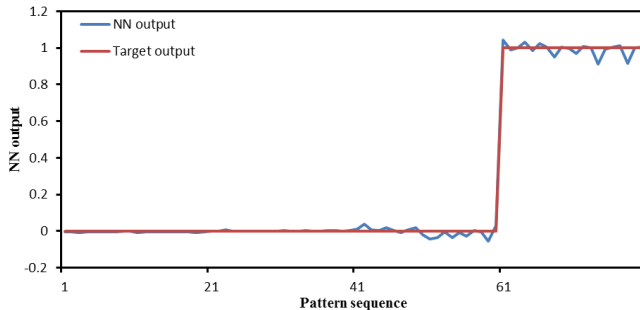


b) Inner race bearing fault

Figure 13a. The designed ANN network performance testing with unseen healthy and faulty data



c) 1mm hole in the outer race



d) 2mm hole in the outer race

Figure 13b. The designed ANN network performance testing with unseen healthy and faulty data

9. CONCLUSION

The development of the intelligent fault diagnosis system has been presented in this paper. Literature concerning industrial robot fault detection using different techniques was first reviewed. Joint three of the PUMA 560 robot was selected to simulate different bearing faults. The mechanical construction of this joint was comprehensively assessed to establish its assembly/disassembly. A discussion about different types of faults that may appear in the joint's gears and bearings is also included, and the physical simulation of several faults, with different severities, in one of bearings of the selected joint were presented. However, to design a reliable monitoring system, appropriate signal analysis techniques have to be employed, since these have a significant impact on the sensitivity of the features extracted from the signals captured. Wavelet analysis represents an efficient method for non-stationary signal analysis and therefore was used here. In the experimental work three single axis, MEMS type accelerometers were utilized to measure the robot vibration in three axes (X, Y and Z). The optimum location that can be used to affix the accelerometers on the robot was justified and found to be close to the robot wrist. Hybrid programming combining LabVIEW graphical programming with Matlab textural programming has been shown to be an effective method to build a signal monitoring and feature extraction system. To design an appropriate ANN and then evaluate its performance for the fault diagnosis, code was developed based on Matlab neural network toolbox. The ANN was trained to distinguish among different types of faults in the robot. A significant level of accuracy in fault diagnosis the ANN has been obtained and the percentage of correct classification was approximately 100%. The logical direction to further extend the capabilities of the designed system is to develop it to diagnosis other categories of faults that are related to joint motor and gears. Also, the examined faults scenarios were simple, as only one fault is seeded at a time; however complex scenarios could be studied by combining multiple fault types at the same time.

REFERENCES

- ABDUL, S. & LIU, G. Decentralised fault tolerance and fault detection of modular and reconfigurable robots with joint torque sensing. *Proceedings - IEEE International Conference on Robotics and Automation*, 2008. 3520-3526.
- ABDULHADY JABER, A. & BICKER, R. 2016. Industrial Robot Backlash Fault Diagnosis Based on Discrete Wavelet Transform and Artificial Neural Network. *American Journal of Mechanical Engineering*, 4, 21-31.
- AL-BADOUR, F., SUNAR, M. & CHEDED, L. 2011. Vibration analysis of rotating machinery using time-frequency analysis and wavelet techniques.

- Mechanical Systems and Signal Processing*, 25, 2083-2101.
- ANIL KUMAR, T. C., SINGH, G. & NAIKAN, V. N. A. 2015. Effectiveness of vibration monitoring in the health assessment of induction motor. *International Journal of Prognostics and Health Management*, 6.
- BOUKABACHE, H., ESCRIBA, C., ZEDEK, S. & FOURNIOLS, J. Y. 2013. Wavelet decomposition based diagnostic for structural health monitoring on metallic aircrafts: Case of crack triangulation and corrosion detection. *International Journal of Prognostics and Health Management*, 4.
- BRITISH-STANDARD 2007. Gears — Cylindrical involute gears and gear pairs — Concepts and geometry.
- DATTA, A., MAVROIDIS, C., KRISHNASAMY, J. & HOSEK, M. Neural network based fault diagnostics of industrial robots using wavelet multi-resolution analysis. American Control Conference, 2007 USA. 1858-1863.
- DEBDAS, S., M.F.QUERESHI, A.REDDY, D.CHANDRAKAR & D.PANSARI 2011. A Wavelet based multiresolution analysis for real time condition monitoring of AC machine using vibration analysis. *International Journal of Scientific and Engineering Research*, 2.
- FILARETOV, V. F., VUKOBRATOVIC, M. K. & ZHIRABOK, A. N. 1999. Observer-based fault diagnosis in manipulation robots. *Mechatronics*, 9, 929-939.
- GHODS, A. & LEE, H. H. A frequency-based approach to detect bearing faults in induction motors using discrete wavelet transform. Proceedings of the IEEE International Conference on Industrial Technology, 2014. 121-125.
- HALME, J. Condition monitoring of a material handling industrial robot. 19th International Congress, 2006 Lulea, Sweden.
- HEATON, J. 2008. *The Number of Hidden Layers* [Online]. <http://www.heatonresearch.com/node/707>.
- JABER, A. & BICKER, R. Industrial Robot Fault Detection Based on Wavelet Transform and LabVIEW. IEEE First International Conference on Systems Informatics, Modelling and Simulation, 2014 Sheffield, United Kingdom. IEEE Computer society.
- JABER, A. A. & BICKER, R. 2015. Real-Time Wavelet Analysis of a Vibration Signal Based on Arduino-UNO and LabVIEW. *International Journal of Materials Science and Engineering*, 3, 66-70.
- JABER, A. A. & BICKER, R. 2016. Fault diagnosis of industrial robot gears based on discrete wavelet transform and artificial neural network. *Insight: Non-Destructive Testing and Condition Monitoring*, 58, 179-186.
- LIANG, X., ZUO, M. J. & HOSEINI, M. R. 2015. Vibration signal modeling of a planetary gear set for tooth crack detection. *Engineering Failure Analysis*, 48, 185-200.
- LIM, W. L. 2009. *The application of artificial neural networks for sensor validation in diesel engine condition monitoring and fault diagnosis*. M. Phil., University of Newcastle upon Tyne.
- MAZUMDAR, J. 2006. *SYSTEM AND METHOD FOR DETERMINING HARMONIC CONTRIBUTIONS FROM NONLINEAR LOADS IN POWER SYSTEMS*. Ph.D Thesis, Georgia Institute of Technology.
- MISITI, M., MISITI, Y., OPPENHEIM, G. & POGGI, J.-M. 1997. *Wavelet Toolbox For Use with MATLAB*, MathWorks.
- MISITI, M., MISITI, Y., OPPENHEIM, G. & POGGI, J.-M. 2001. *Wavelet Toolbox for Use With Matlab*, MathWorks.
- MOHANTY, A. R. 2015. *MACHINERYCONDITION MONITORING: PRINCIPLES AND PRACTICES*, Taylor & Francis Group.
- NEGNEVITSKY, M. 2005. *Artificial intelligence: a guide to intelligent systems*, ADDISON WESLEY.
- OLSSON, E., FUNK, P. & XIONG, N. 2004. Fault diagnosis in industry using sensor readings and case-based reasoning. *Journal of Intelligent and Fuzzy Systems*, 15, 41-46.
- PAN, M. C., VAN BRUSSEL, H. & SAS, P. 1998. Intelligent joint fault diagnosis of industrial robots. *Mechanical Systems and Signal Processing*, 12, 571-588.
- PANDYA, D., UPADHYAY, S. & HARSHA, S. 2012. Ann based fault diagnosis of rolling element bearing using time-frequency domain feature. *International Journal of Engineering Science and Technology*, 4, 2878-2886.
- RODRIGUEZ-DONATE, C., MORALES-VELAZQUEZ, L., OSORNIO-RIOS, R. A., HERRERA-RUIZ, G. & ROMERO-TRONCOSO, R. J. 2010. FPGA-based fused smart sensor for dynamic and vibration parameter extraction in industrial robot links. *Sensors*, 10, 4114-4129.
- S.N.SIVANANDAM., S.SUMATHI. & S.N.DEEPA. 2006. *Introduction to Neural Networks Using MATLAB 6.0*, McGraw Hill.
- SAWICKI, J. T., SEN, A. K. & LITAK, G. 2009. Multiresolution wavelet analysis of the dynamics of a cracked rotor. *International Journal of Rotating Machinery*, 2009.
- SPONG, M. W., HUTCHINSON, S. & VIDYASAGAR, M. 2005. *Robot Modeling and Control*, Wiley.
- SUBBARAJ, P. & KANNAPIRAN, B. 2014. Fault detection and diagnosis of pneumatic valve using Adaptive Neuro-Fuzzy Inference System approach. *Applied Soft Computing Journal*, 19, 362-371.

- VAN, M., KANG, H. J. & RO, Y. S. 2011. A robust fault detection and isolation scheme for robot manipulators based on neural networks.
- VIVAS, E. L. A., GARCIA-GONZALEZ, A., FIGUEROA, I. & FUENTES, R. Q. Discrete Wavelet transform and ANFIS classifier for Brain-Machine Interface based on EEG. 2013 6th International Conference on Human System Interactions, HSI 2013, 2013. 137-144.
- YILDIRIM, A. & ESKI, I. 2010. Noise analysis of robot manipulator using neural networks. *Robotics and Computer-Integrated Manufacturing*, 26, 282-290.
- YUAN, J., LIU, G. & WU, B. 2011. Power efficiency estimation-based health monitoring and fault detection of modular and reconfigurable robot. *IEEE Transactions on Industrial Electronics*, 58, 4880-4887.