BEARING FAULTS DETECTION USING DEEP LEARNING MODELS

A report submitted for

ICIR16- SUMMER INTERNSHIP

By

ABHAY KUMAR GUPTA

(110119004)

DEPARTMENT OF INSTRUMENTATION AND CONTROL ENGINEERING

NATIONAL INSTITUTE OF TECHNOLOGY

TIRUCHIRAPALLI – 6200015

INTERNSHIP CERTIFICATE

Department of Computer Applications

NATIONAL INSTITUTE OF TECHNOLOGY

(Formerly Regional Engineering College)

Tiruchirappalli – 620015, Tamil Nadu, India

PHONE: +91 (431) 2503746; +91 8438207784; E -MAIL: usreddy@nitt.edu

Dr. U. Srinivasulu Reddy
Associate Professor

Sub: Internship at National Institute of Technology, Tiruchirappalli.

I am happy to inform that **Mr. ABHAY KUMAR GUPTA** (Reg. no. 110119004) currently pursuing Instrumentation and Control Engineering at National Institute of Technology, Tiruchirappalli- 620015 has completed his internship entitled "**Bearing Faults detection using Deep Learning models**" successfully in offline mode under my guidance for a period of nine weeks i.e. 18th May, 2022 to 31st July, 2022. He has done his best and successfully executed the given tasks. I wish him success in all his endeavours.

(Dr. U. Srinivasulu Reddy)

Dr. U. Srinivasulu Reddy, MCA, M.Phil(CS), PhD.
Associate Professor
Department of Computer Applications
National Institute of Technology
Tiruchirappalli - 620 015. Tamil Nadu, India

TABLE OF CONTENTS

Title **Page No.**

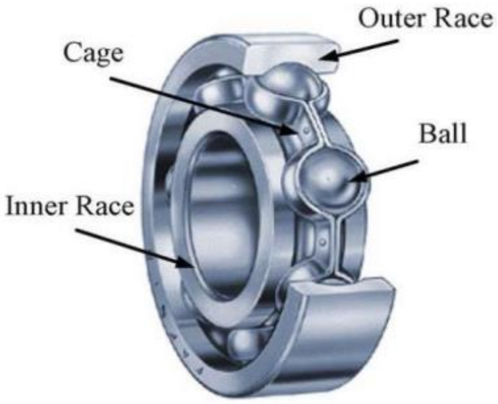
1.ABSTRACT

In rotating machinery, rolling element bearings (REBs) are prone to failure, which could result in significant production losses and create a threat to human safety. Bearing fault diagnostic (BFD) is the process of detecting, isolating, and identifying defects before they cause a failure. The classic methods of fault classification require system experts with specific domain expertise to identify the discriminating fault features using signal processing techniques. These approaches are person-dependent, arduous, time-consuming, and prone to errors. It is crucial to find REB faults early on when they are minor and can be fixed automatically. Using sensor-acquired vibration data from rotary machine operations, Deep Learning (DL) models like Convolution Neural Network (CNN) can detect and classify REB faults.

2. INTRODUCTION

Electric machines are widely employed in a variety of industry applications and electrified transportation systems. For certain applications these machines may operate under unfavourable conditions, such as high ambient temperature, high moisture, and overload, which can eventually result in motor malfunctions that lead to high maintenance costs, severe financial losses, and safety hazards. The malfunction of electric machines can be generally attributed to various faults of different categories, including drive inverter failures, stator winding insulation breakdown, bearing faults and air gap eccentricity. Several surveys regarding the likelihood of induction machine failures conducted by the IEEE Industry Application Society (IEEE-IAS) and the Japan Electrical Manufacturers' Association (JEMA) reveal that bearing fault is the most common fault type and is responsible for 30% to 40% of all the machine failures.

The structure of a rolling-element bearing is illustrated in Fig. 1, which contains the outer race typically mounted on the motor cap, the inner race to hold the motor shaft, the balls or the rolling elements, and the cage for restraining the relative distances between adjacent rolling elements. The four common scenarios of misalignment that are likely to cause bearing failures are demonstrated in Fig. 1(a) to (d). Since bearing is the most vulnerable component in a motor drive system, accurate bearing fault diagnostics has been a research frontier for engineers and scientists for the past decades. Specifically, this problem has been approached by developing a physical model of bearing faults and understanding the relationship between bearing faults and measurable signals, which can be captured by a variety of sensors and analysed with signal processing techniques.



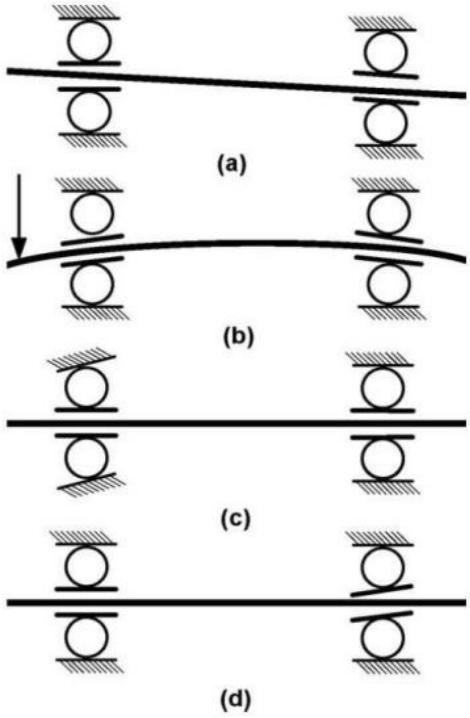


Fig 1

In this context, this report seeks to present a thorough overview of the recent research work devoted to applying ML and DL techniques on bearing fault diagnostics. In Section III, we introduce some of the most popular datasets used for bearing fault detection.

Next, in Section IV we investigate some traditional ML methods, including PCA, k-nearest neighbors (k-NN), SVM, etc., with a brief overview of major publications applying each ML algorithm for bearing fault detection. For the main part of this paper, in Section IV, we take a deep dive into the research frontier of DL based bearing fault identification. In this section, we will provide our understanding of the research trend toward DL approaches. Specifically, we will discuss the advantages of DL based methods over the conventional ML methods in terms of fault feature extraction and classifier performance, as well as new functionalities offered by DL techniques that cannot be accomplished before. We will also provide a detailed analysis of each of the major DL techniques, including Artificial neural network (ANN), 1-D convolutional neural network (1D-CNN), 2-D convolutional neural network (2D-CNN).

3. DATASET DESCRIPTION

Data is the foundation for all of the ML methods. To develop effective ML and DL algorithms for bearing fault detection, a good collection of datasets is necessary. Since the natural bearing degradation is a gradual process and may take many years, most people conduct experiments and collect data either using bearings with artificially induced faults, or with accelerated life testing methods. While the data collection is still time consuming, fortunately a few organizations have made the effort and published their bearing fault datasets for engineers and researchers to develop their own ML algorithms. Thanks to their prevalence in the research community, these datasets can also serve as a common ground for the evaluation and comparison of different algorithms. Before getting into details of various ML and DL developments, in this section, we briefly introduce a popular database used by most researchers for the Bearing Fault Diagnosis.

3.1CASE WESTERN RESERVE UNIVERSITY (CWRU) DATASET:

The Case Western Reserve University (CWRU) conducted experiments for collecting the REB data for normal and faulty bearings, using a Reliance Electric motor driving a shaft on which, a torque transducer and encoder are mounted. Torque was applied to the shaft via a dynamometer and electronic control system. There are two bearings available in the setup, one in the drive-end (DE) and the other in the fan-end (FE). Drive end bearing is 6205-2RS JEM, which is of SKF make deep groove ball bearing, and this bearing alone was considered for our proposed work. Accelerometers were placed at the 12 o'clock position at the DE of the motor housing, as well as in the FE. The acceleration data from both ends (DE and FE), were acquired. CWRU dataset was used for training and validation of the proposed 2D-CNN architecture. The experimental CWRU bearing set-up is shown in Fig. 2. The CWRU dataset contains the recorded REB data for normal and faulty bearings. Motor bearings' data with seeded single faults using Electro-Discharge

Machining (EDM) for the fault sizes of 0.007 inches, 0.014 inches and 0.021 inches in diameter at the inner race fault (IRF), ball fault (BLF) and outer race fault (ORF) respectively, were considered for the proposed model. The CWRU dataset serves as a fundamental dataset to validate the performance of Different ML and DL Algorithms.

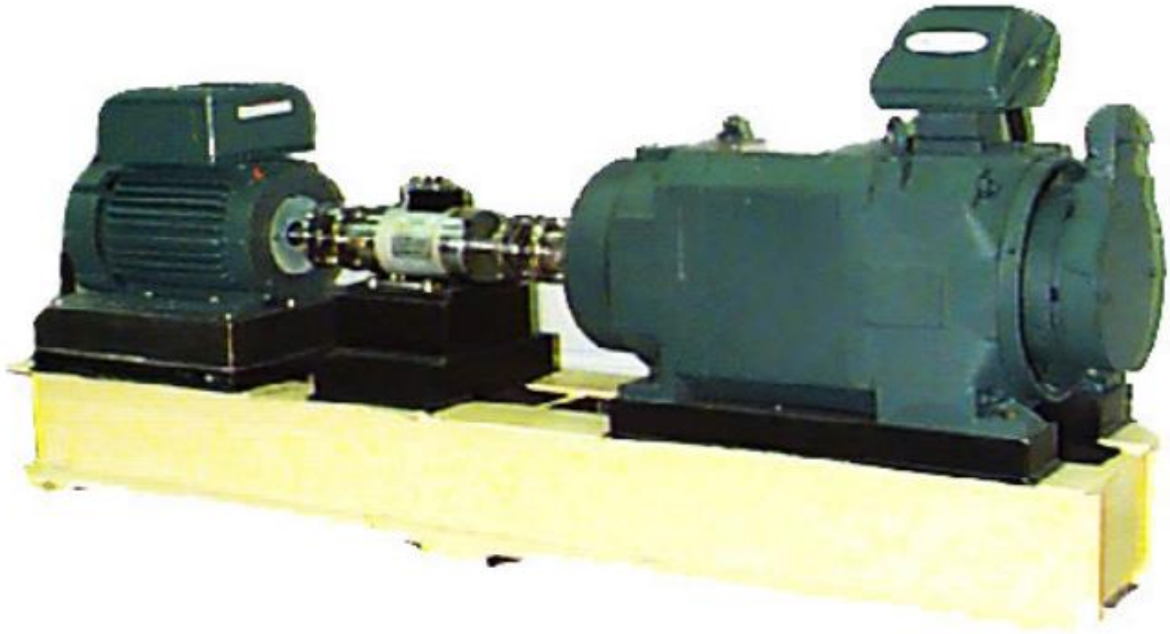


Fig 2

3.2 DATA PREPOCESSING:

- Libraries that are imported:

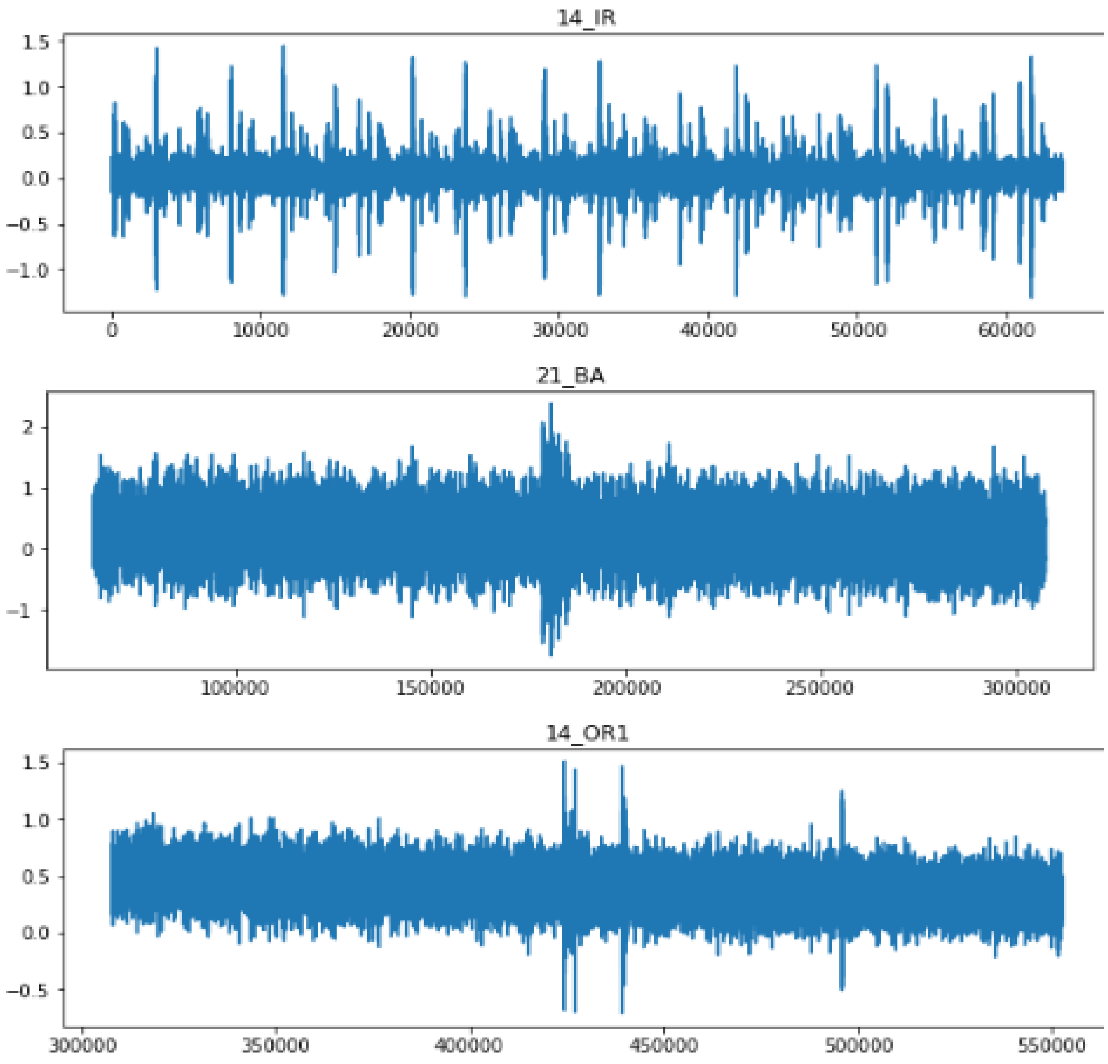
  Scipy.io, NumPy, pandas, os, matplotlib. pyplot

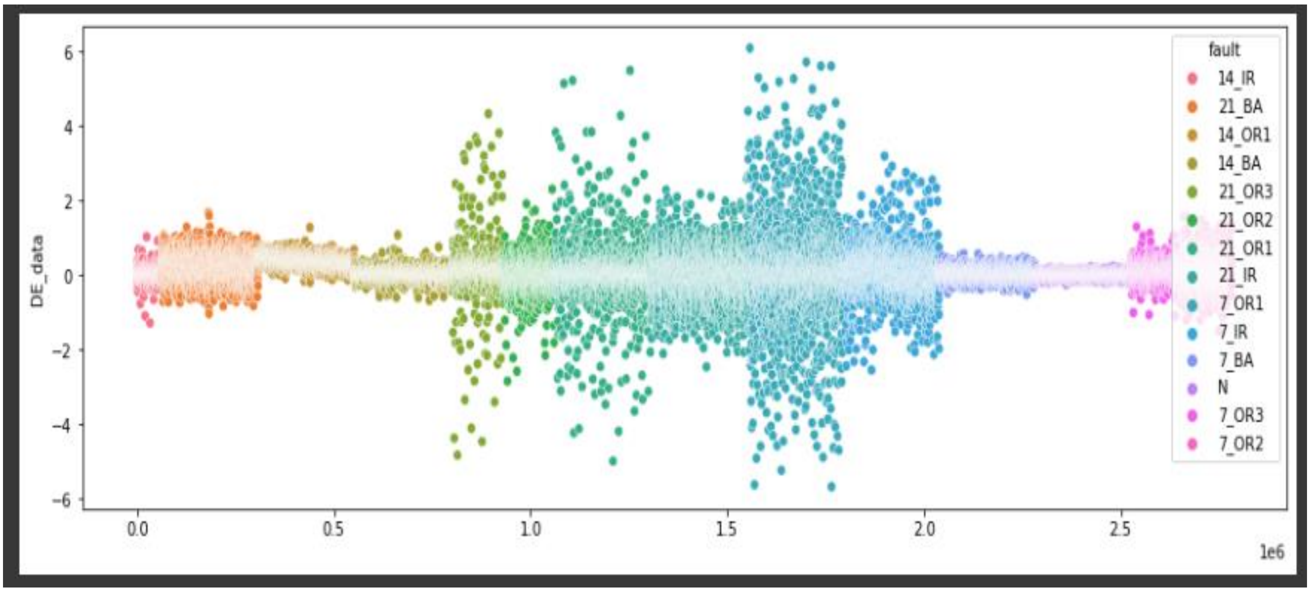
- After importing the libraries and loading the dataset, the root directory is joined to a variable 'file_name' for each of the '.mat ' files.

- The data in mat files are being stored in dictionary, so the 'keys' of the dictionary will give the columns in the dataset.

- In all the columns only ‘drive_end” data is being considered and rest are omitted because only drive_end contribute in detecting the type of faults.
- After getting the ‘drive_end’ data, .mat file is converted to .csv file, so Machine Learning and Deep Learning models can be applied easily.

3.3 DATA VISUALISATION:





For each of the faults, once the data is plotted separately. Then the combined plot for all the faults is plotted.

4. MACHINE LEARNING MODELS

To detect the presence of a bearing fault using a classical ML algorithm, the characteristic fault frequencies are calculated based on the rotor mechanical speed and the specific bearing geometry, and these frequencies will serve as fault features. This feature determination process is known as ‘‘feature engineering’’. The amplitude of signals at these frequencies can be monitored to train various ML algorithms and identify any anomalies. However, such a technique may encounter many challenges that ultimately affect the classification accuracy.

1) Sliding: The fault frequency assumes that no sliding occurs between the rolling element and the bearing raceway, i.e., these rolling elements will only roll on the raceway. Nevertheless, this is seldom the case, as the rolling element often undergoes a combination of rolling and sliding movement. Therefore, the calculated frequency may deviate from the real fault frequency and make this manually determined feature less informative of a bearing defect.

2) Frequency Interplay: If multiple types of bearing faults occur simultaneously, these faults will interact and the resultant characteristic frequencies can add or subtract due to a complicated electro-mechanical process, thereby obfuscating the informative frequencies.

3) External Vibration: There is also the possibility of interference induced from additional sources of vibration, i.e. bearing looseness and environment vibration, which can obscure the useful features.

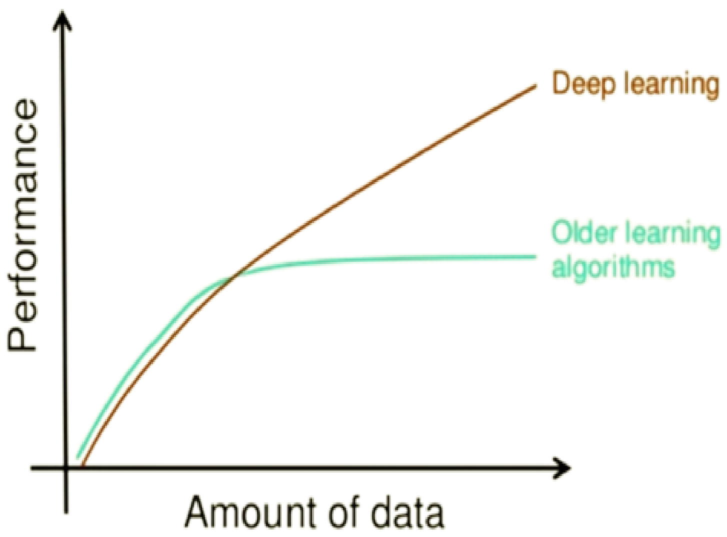
4) Observability: Some faults, such as the bearing lubrication and general roughness related faults, do not even manifest themselves as a characteristic cyclic frequency, which makes them very hard to detect with the traditional model-based spectral analysis or classical data-driven ML methods.

5) Sensitivity: The sensitivity of various features that are characteristic of bearing defects may vary considerably at different operating conditions. A very thorough and systematic ‘‘learning stage’’ is typically required to test the sensitivity of these frequencies on any desirable operating condition before it can be actually put into use with the traditional approach.

Because of the challenges, manually engineered features based on the bearing characteristic fault frequency can be difficult to interpret, and sometimes may even lead to inaccurate classification results, especially when applying the ‘‘shallow’’ classical ML methods that rely on human-engineered features in the training process. Therefore, many DL algorithms with automated feature extraction capabilities and better classification performance have been applied to bearing fault diagnostics, which will be discussed in detail in the next section.

5. NEED FOR DEEP LEARNING

- Nowadays, Deep Learning is the most attractive research trend in Machine Learning. With the ability of learning features from raw data by deep architectures with many layers of non-linear data processing units, Deep Learning has become a promising tool for intelligent bearing fault diagnosis.



- A big advantage of using deep learning models, and a key part in understanding why it's important for classification, is that it's powered by massive amounts of data.

- When there is lack of domain understanding for feature introspection, Deep Learning techniques outshine others as you worry less about feature engineering

6. DEEP LEARNING MODELS

Deep learning is a subset of machine learning that achieves great power and flexibility by learning to represent the world as nested hierarchy of concepts, with each concept defined in relation to simpler concepts, and more abstract representations computed from less abstract ones. On smaller datasets, classical ML algorithms can compete with or even outperform deep learning networks. With the increase of the amount of data, the performance of DL can significantly outperform most classical ML algorithms.

The complexity of the computed function grows exponentially with model depth. DL has the best-in-class performance that significantly outperforms other solutions to problems in multiple domains, including speech, language, vision, game playing, etc. DL also removes the need for feature engineering. One can simply pass the data directly to the neural network, and the network can automatically learn the features from raw data by auto tuning the weights in the network. The DL network eliminates completely the challenging stage of feature engineering.

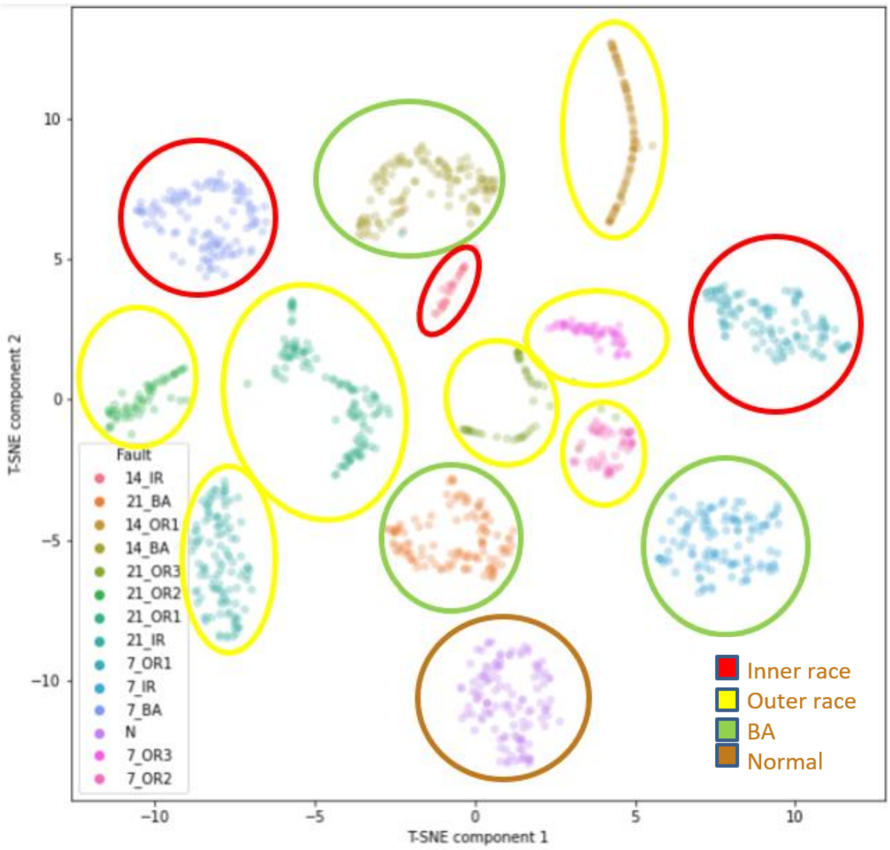
6.1 ANN MODEL IMPLEMENTATION:

- In the literature, the use of neural networks for fault diagnosis is of great interest. This intense focus is due to the ability of ANNs to learn complex nonlinear relationships from a set of training examples.
- An ANN consists of a set of simple processing units that model the biological neurons of the human brain. Each artificial neuron is an elementary processor. It accepts a series of entries (In_j) weighted by synaptic coefficients (W_{ij}) and generates an output (out_i), generally as:

$$out_i = f\left(\sum_{j=1}^{n} W_{ij} In_j + b_i\right)$$

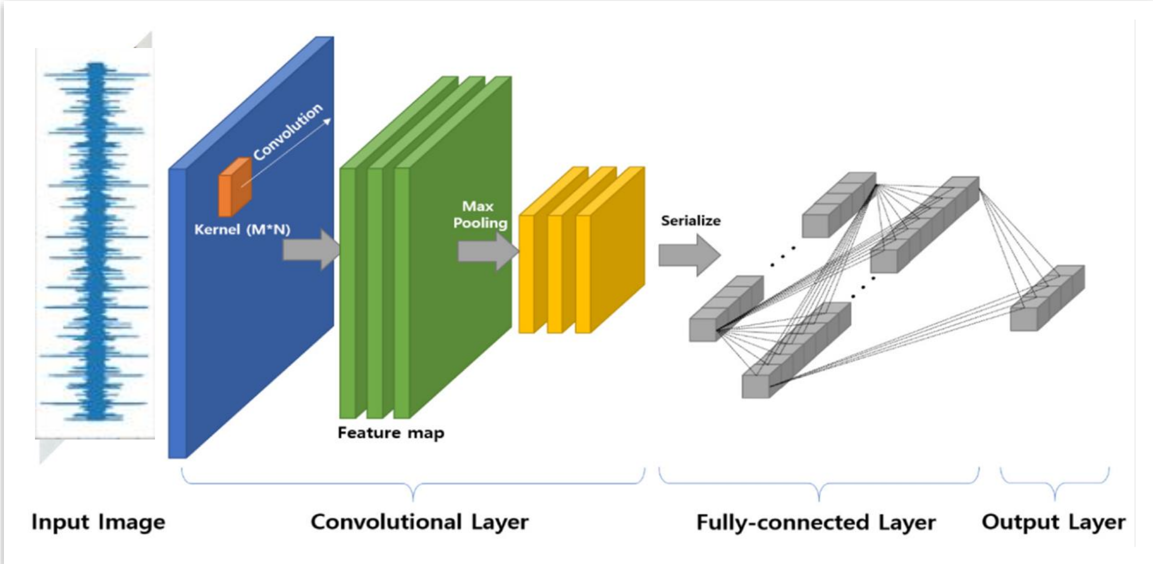
- Where i=1, . . ,m
- f is activation function
- b_i is bias of i^{th} neuron.
- Wij is synaptic coefficients

ANN Model Visualization .



6.2 CNN MODEL IMPLEMENTATION:

- The CNN, one of the most representative deep learning algorithms, combines deep structure and convolution calculation.
- It is widely utilized in the field of image processing. It can continuously abstract these properties by automatically extracting them from the input data layer by layer.



- Accurate classification outcomes can be achieved using the extracted abstract characteristics.
- Convolutional, pooling, and fully linked layer structures are the three main layer structures in a CNN.
- The convolutional layer and the pooling layer are put together to form the convolution module, which is then used to extract the features from the input data using a deep network structure made up of several convolution modules.
- The fully connected layer is the last layer of the network model, and it is used to perform the classification task.

6.3 1D CNN MODEL:

- In 1D-CNN model, the CNN is established to directly process the 1-D vibration signals.
- Reasons to approach 1D CNN based model:
    1. Its compact architecture configuration (rather than the complex deep architectures) which performs only 1D convolutions making it suitable for real-time fault detection and monitoring.
    2. Its cost effective and practical real-time hardware implementation,
    3. Its ability to work without any pre-determined transformation (such as FFT or DWT), hand-crafted feature extraction and feature selection.
    4. Its capability to provide efficient training of the classifier with limited size of training data set and limited number of BP iterations.
- Fault diagnosis can be much simpler if the model parameter is directly obtained from the physical coefficients. Time-domain frequency-domain and time–frequency analysis techniques have been applied as signal-based systems for extracting the health information from the measured data.
- In this we use the window length and stride (500,300).

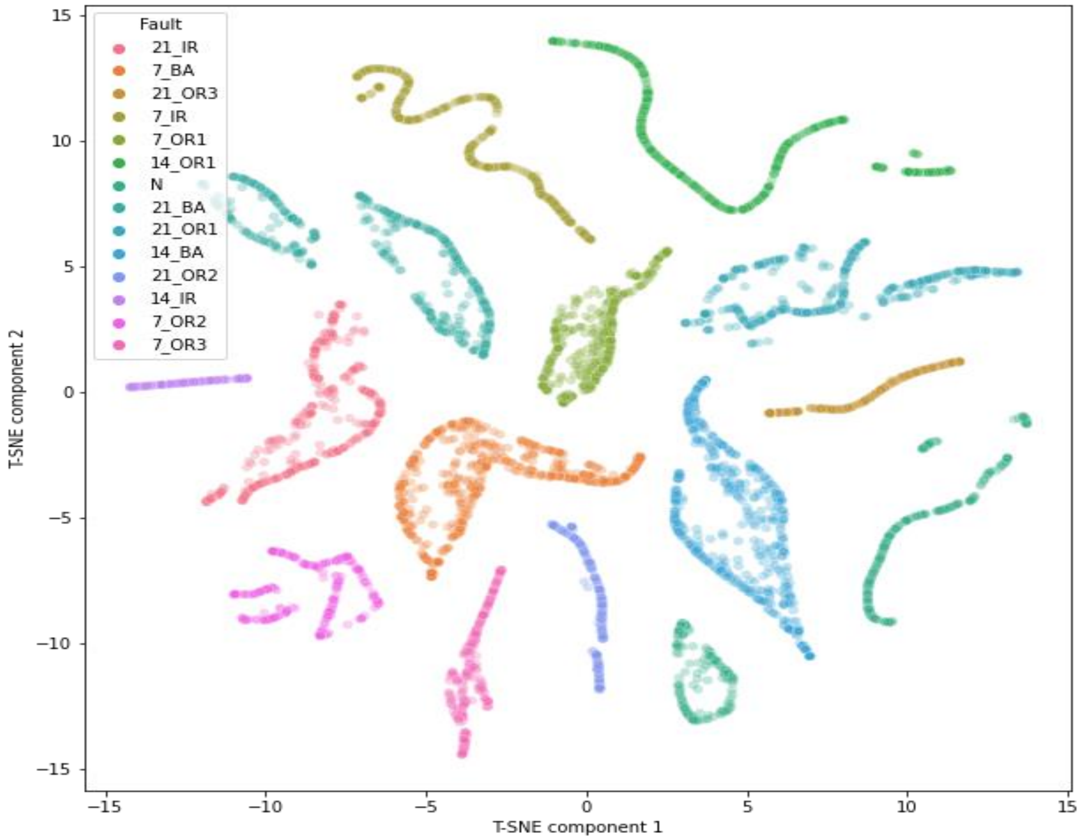
1D CNN MODEL with Multiple Kernels:

- Here we use different lengths of kernels to capture various frequency features and their performance evaluated. Outputs of hidden layers are visualized using t-SNE dimensionality reduction.
- Highest accuracy :

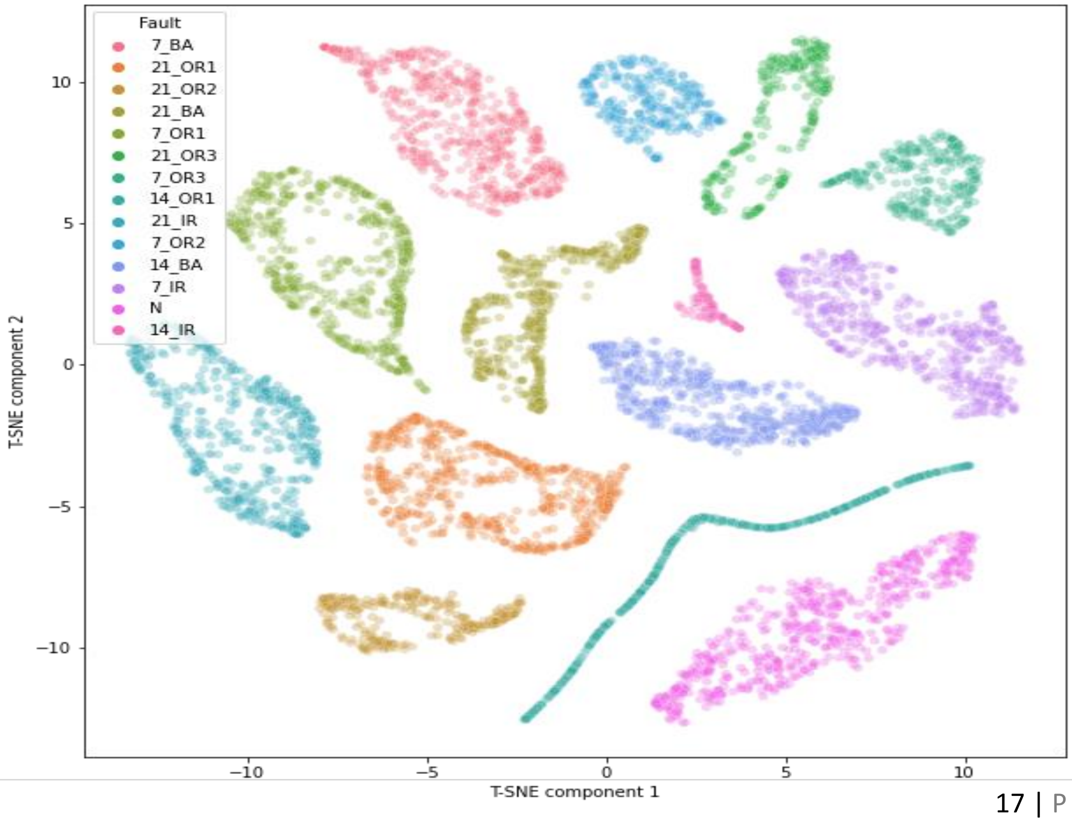
    For 1D CNN Bearing vibration : 1.00

    For 1D CNN with multiple Kernels : 1.00

1D CNN Model Visualization:

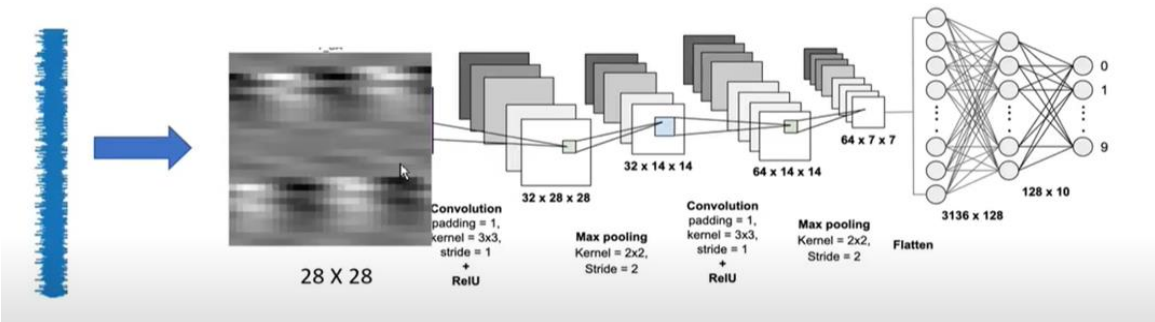


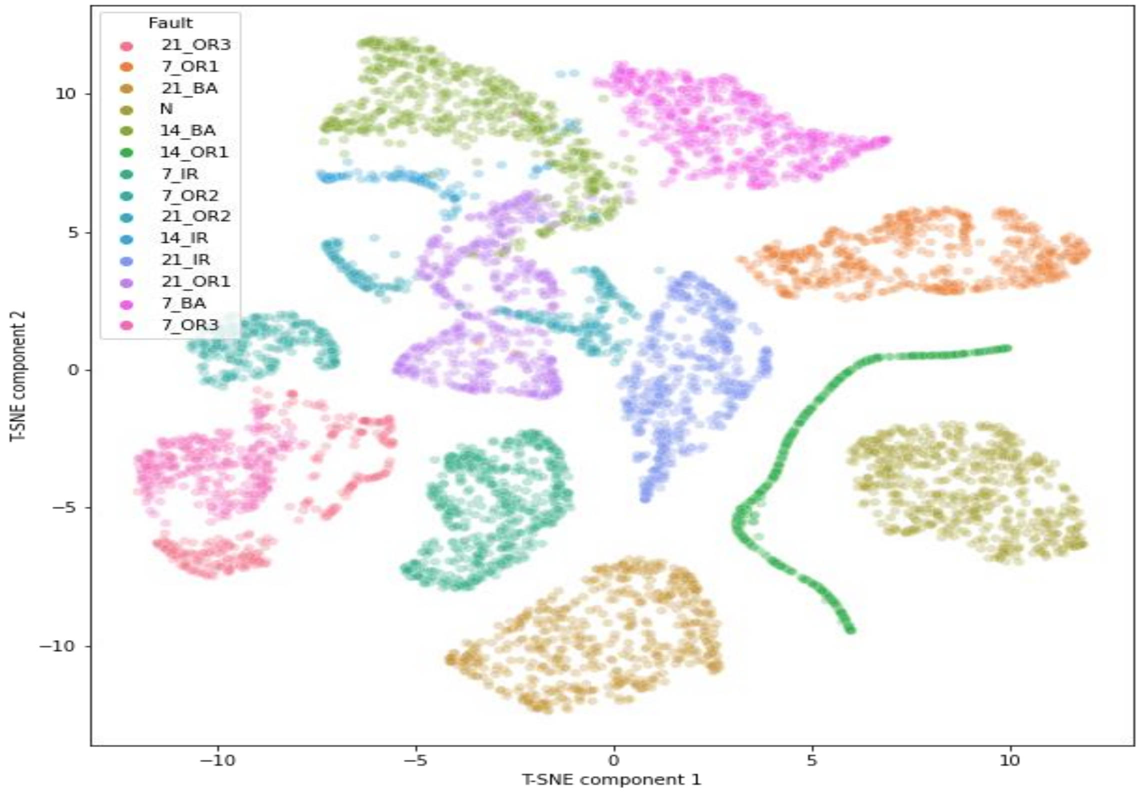
1D CNN Model with Multiple Kernels Visualization:



6.4 2D-CNN MODEL

- Whereas the 2D-CNN, which is the intelligent diagnosis model, is more suitable for extracting the feature information from the 2-D image. Therefore, the 2-D preprocessing stage is set to convert the 1-D vibration signal into the 2-D gray image.
- A window length of 784 (28*28) is taken from the vibration signal to visualize it in a grayscale image.
- The grayscale images thus formed are given as the input for the convolutional operation.



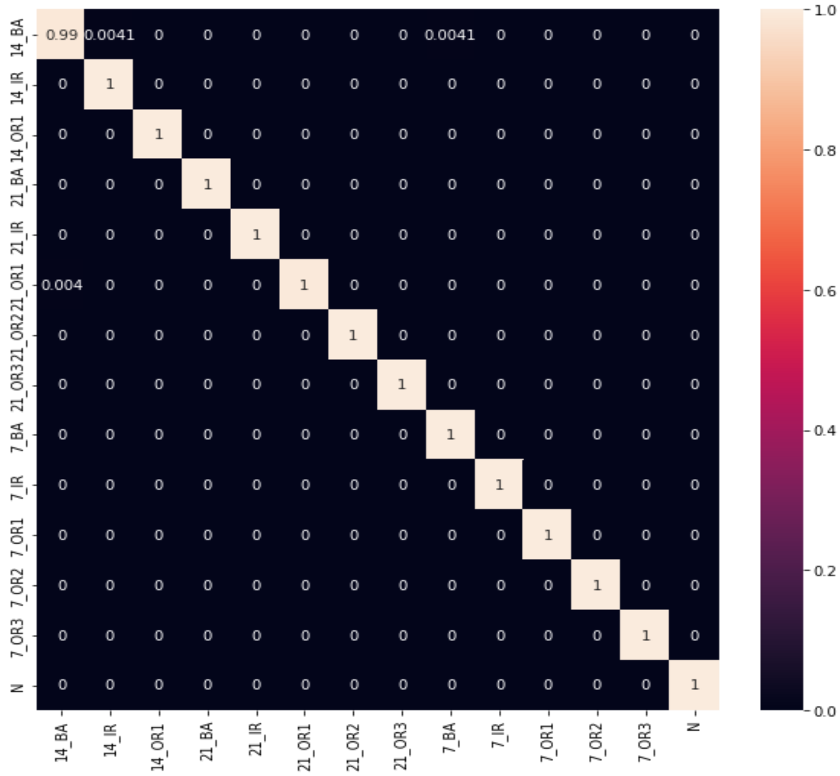


7. RESULTS:

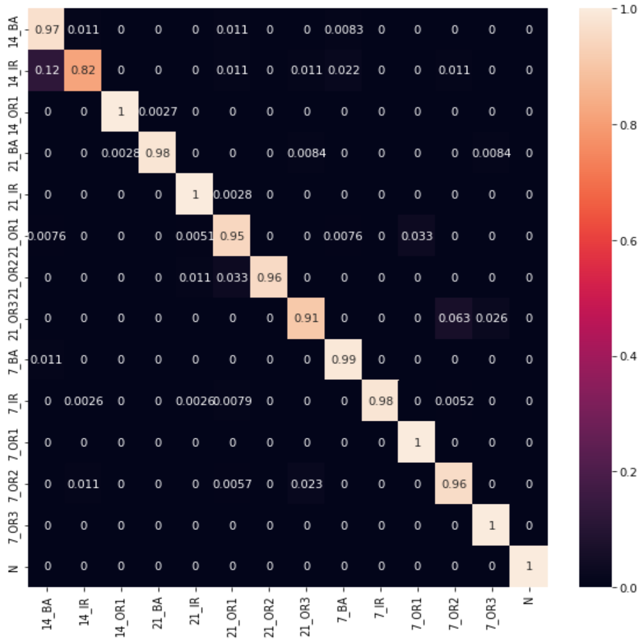
The proposed 2D-CNN network was trained with the accelerometer data acquired from both ends (DE and FE). The DE and FE data available was sampled at sampling rates of 48K samples per second (KS/s). In the proposed work, a total of 14 classes were considered for the fault classification. It includes normal, three number of faults(0.007, 0.014, 0.021 inches) of IRF, BLF and ORF respectively. Further, three positions are considered for the ORF (analogous to 3, 6, and 12 o'clock positions). The data available at the CWRU repository have different load conditions ranging from 1 HP to 3 HP, and also at four different speeds 1797, 1772, 1750, and 1730 rotations per minute (RPMs). Even with different load conditions and speeds of the motor, the fault classification was not affected. The entire CWRU dataset with all the aforementioned diversities were considered for training the proposed DL models.

The Confusion matrices of ANN, 1D-CNN and 2D-CNN are shown in the below figures.

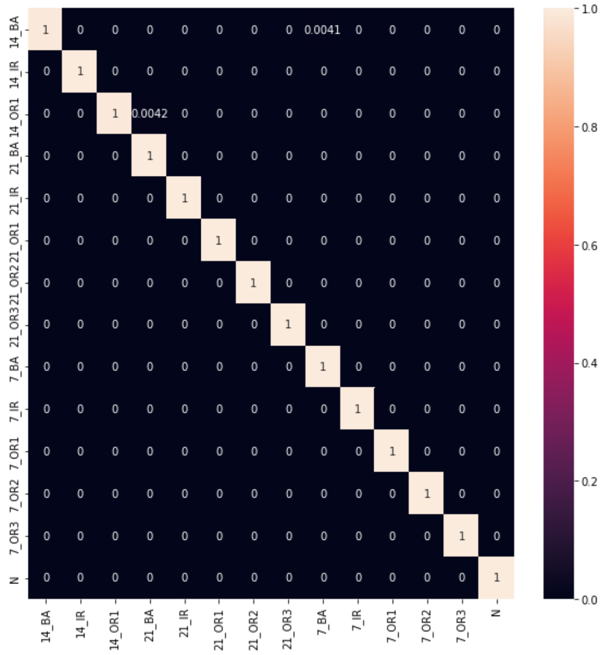
ANN MODEL



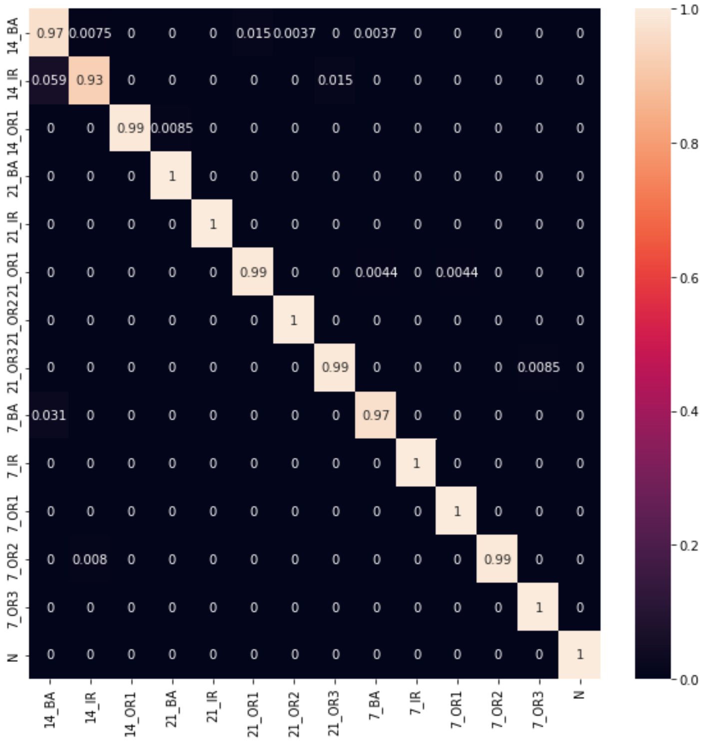
1D CNN MODEL



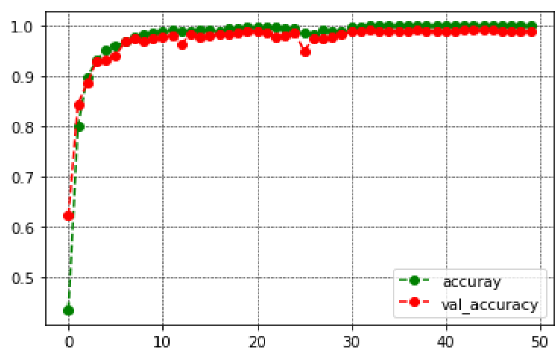
1D CNN With Multiple Kernels



2D CNN Classification



Plot of accuracy and Val. accuracy



7.1 CONCLUSION

The proposed 2D-CNN architecture was trained and tested with the diversity of the CWRU bearing dataset (Accelerometers' data acquired from both DE and FE, different sampling rates namely 12 Khz and 48 Khz, three load conditions from 1 HP to 3 HP, Four different speeds, and ORF faults in 3, 6, and 12'o clock positions), for the generalization and faster convergence with best accuracy. The 2D-CNN model considered all the diversities of CWRU dataset the model and was able to achieve 99.74% testing accuracy with 20 epochs itself. The generalization capability of the proposed 2D-CNN architecture makes it a suitable model for real-world applications.

7.2 REFERENCES

1. https://papers.phmsociety.org/index.php/phme/article/download/2947/1761
2. https://www.mdpi.com/1424-8220/17/3/549/pdf-vor
3. https://www.kaggle.com/code/dlagazo/ims-bearing-dataset
4. https://www.kaggle.com/brjapon/cwru-bearing-datasets
5. https://link.springer.com/article/10.1007/s11831-018-9286-z
6. https://stats.stackexchange.com/questions/371333/is-it-important-to-make-a-feature-scaling-before-using-gaussian-mixture-model
7. http://mkalikatzarakis.eu/wp-content/uploads/2018/12/IMS_dset.html
8. https://sequentia.readthedocs.io/en/latest/sections/classifiers/gmmhmm.html
9. https://hmmlearn.readthedocs.io/en/latest/tutorial.html#building-hmm-and-generating-samples
10. https://github.com/eonu/sequentia/blob/master/lib/sequentia/classifiers/hmm/gmmhmm.py
11. https://www.mathworks.com/matlabcentral/answers/351766-how-to-interpret-log-probability-from-hmmdecode
12. https://www.researchgate.net/publication/224177188_A_Mixture_of_Gaussians_Hidden_Markov_Model_for_failure_diagnostic_and_prognostic
13. https://www.kaggle.com/code/yasirabd/rul-nasa-bearing-mean-rms-skewness-kurtosis
14. https://www.kaggle.com/code/yasirabd/nasa-bearing-feature-extraction/notebook
15. https://www.kaggle.com/code/furkancitil/nasa-bearing-dataset-supervised-learning