ISSN: 18158129 E-ISSN: 18151027
Volume 18, Issue 08, August, 2022

Prediction and early fault detection in agricultural processing lines

Nguyen Le Dung[1,2*], Phan Huynh Lam[1,2]

Faculty of Electrical and Electronics Engineering, Ho Chi Minh City University of Technology (HCMUT), 268 Ly Thuong Kiet Street, District 10, Ho Chi Minh City, Viet Nam[1]
Vietnam National University Ho Chi Minh City, Linh Trung Ward, Thu Duc District, Ho Chi Minh City, Vietnam[2]

Corresponding Author: 1,2*

ABSTRACT— The article presents the application of early prediction to agricultural production lines, a field where technology is rarely applied, and the direction of using artificial intelligence to improve classical algorithms so that forecasting is more flexible and active. Early forecasting will help agricultural factories have maintenance plans, uninterrupted production, and significant economic significance, but current systems have not been applied. The research topic will use 3-axis accelerometer and indirect contact with agricultural equipment for early prediction and fault detection; Does not affect the system, easy to install.

KEYWORDS: Vibration analysis, Predictive maintenance, EFD

1. INTRODUCTION

Vietnam is a long-standing agricultural country and has great agricultural potential on the world map in the fields of rice, cashew, and coconut processing... nowadays, automation mechanisms are gradually replacing humans in the processing field. Agricultural variables and Vietnam applied in most of the large-scale processing factories, leading to maintenance and early detection of defects, are essential to reduce production waste.

Figure I-1 Coconut production line system in Vietnam

Fault detection and prediction is a promising area for the reliability of modern industrial and agricultural systems; the technique of early fault diagnosis (EFD) is attracting more and more attention from academia; this is the strength of artificial intelligence; in this paper, we propose an application to monitor and predict engine failure in agricultural conveyors; where the mechanisms operate frequently, failures often occur in the actuators.

The EFD is essential in providing the proper information to take the necessary maintenance actions and thus prevent serious failures and reduce financial losses. This paper presents three-axis non-contact vibration sensor data collection and value analysis via FFT for fault identification and prediction.

2. Some related studies

This issue is also studied by many firms and researchers, such as:

a) Intel [1] collect the Bearing dataset, and analyze FFT before going through Machine Learning Algorithm sets (Logistics, K-Mean, GMM)

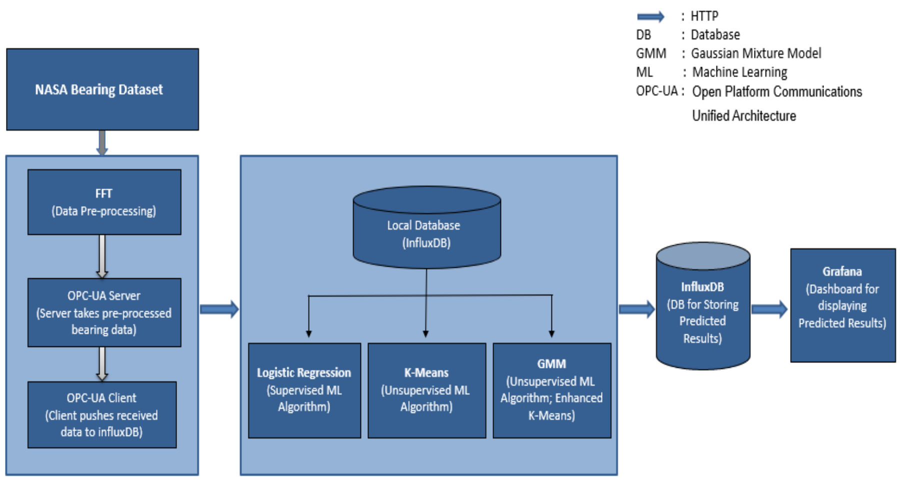
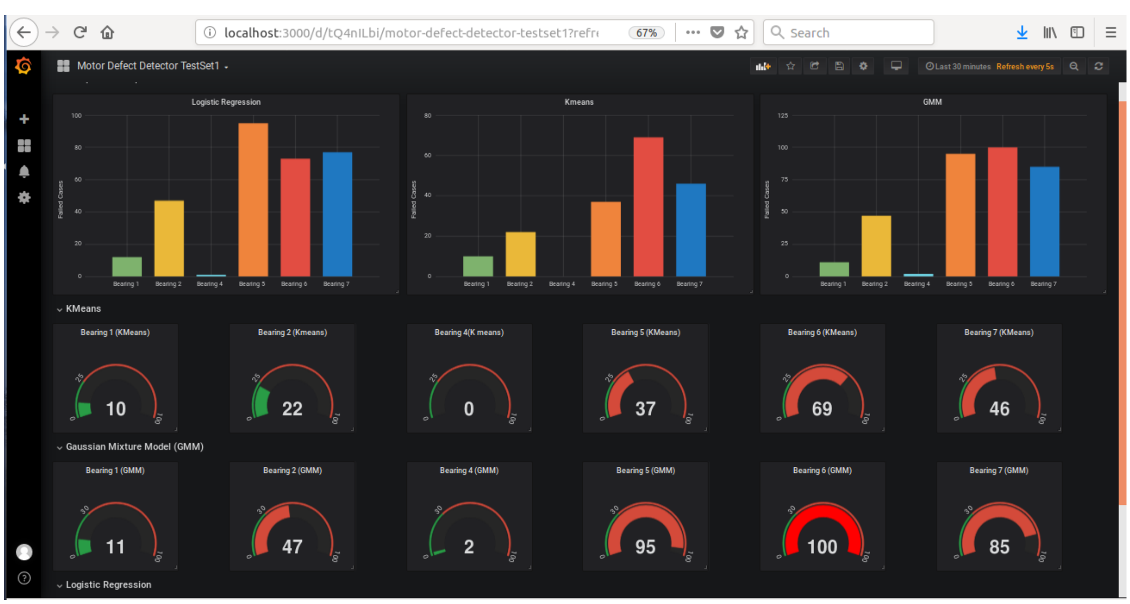


Figure 2. Intel processor block diagram



ISSN: 18158129 E-ISSN: 18151027
Volume 18, Issue 08, August, 2022

Figure 3. Grafana's visual interface.

In this study, Intel shows the basic implementation of FFT, Logistic Regression, K-Means clustering, and GMM. It also shows how helpful the FFT is in Feature Engineering of Oscillating Data of a Machine. Several methods do not require neural network training to detect errors, starting from the most basic (FFT) to the most complex (Gaussian Mixed Model). They have the advantage that they can be reused with minor modifications across different data streams and do not require a lot of previously classified data (unlike neural networks). Some of these methods can be used to classify data to train DNNs. Algorithms used in the method:

FFT: Fast Fourier Transform (FFT) is an algorithm that samples a signal over time (or space) and divides it into frequency components. These components are single sinusoidal oscillations at separate frequencies, each with its own amplitude and phase.

Y = fft (X) computes the discrete Fourier transform (DFT) of X using the fast Fourier transform (FFT) algorithm. If X is a vector, then fft(X) returns the Fourier transform of the vector. More details

Logistic Regression: Logistic regression is a statistical method for analyzing a data set in which one or more independent variables determine an outcome. Outcomes are measured by a dichotomous variable (where there are only two possible outcomes).

In logistic regression, the dependent variable is binary or dichotomous, i.e., it contains only data encoded as 1 (TRUE, success, pregnant, etc.) or 0 (FALSE, failure, no pregnancy, etc.). The goal of logistic regression is to find the best fit (but biologically plausible) model to describe the relationship between the dichotomous characteristic of interest (dependant variable = response variable or outcome variable).) and the set of independent variables (predictor or explanatory). More details

Kmeans Clustering: K-mean clustering is an unsupervised learning style, used when you have unlabeled data (i.e. data with no defined categories or groups). This algorithm aims to find groups in the data, with the number of groups represented by the variable K. The algorithm works iteratively to assign each data point to one of K groups based on the features specified provided. Data points are grouped based on feature similarity.

The result of the K-means clustering algorithm is:
- Centers of K clusters, which can be used to label new data
- Labels for training data (each data point is assigned to a unique cluster)

Each center of a cluster is a set of feature values that define the resulting groups. The centroid feature weight test can be used to explain what kind of cluster each cluster represents qualitatively.

GMM: The Gaussian Mixture Model is a probabilistic model that assumes all data points are generated from a mixture of a finite number of Gaussian distributions with unknown parameters. One can think of mixed models as k-means clustering generalizations to incorporate information about the covariance structure of the data as well as the latent Gaussian centers.

The GaussianMixture object implements an expectation maximization (EM) algorithm to fit Gaussian mixed models. It can also draw confidence ellipsoids for multivariate models and compute Bayesian Information Criteria to evaluate the number of clusters in the data. A GaussianMixture.fit method is provided to learn the Gaussian Mixture Model from the train data. With the test data, it is possible to assign to each Gaussian sample to which it can mainly belong using the GaussianMixture.p Prediction method.

GaussianMixture comes with different options to limit the covariance of different classes to be estimated: spherical, diagonal, forced or full covariance.

b) A Review of Early Fault Diagnosis Approaches and Their Applications in Rotating Machinery (by Yu Wei, Yuqing Li *, Minqiang Xu and Wenhu Huang) [2]

The research team presents an early fault diagnosis method to detect and predict errors. Early fault diagnosis techniques, applications of machine EFD are considered in two aspects: error frequency based method (FFT) and artificial intelligence based method (K-NN, SVM, NN). Algorithms used:

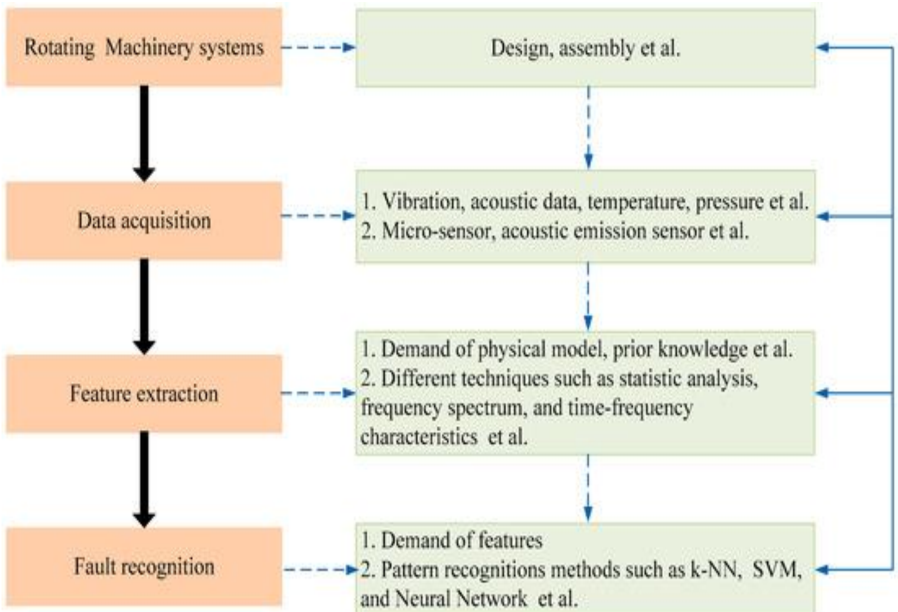


Figure 4. Algorithm flowchart [2]

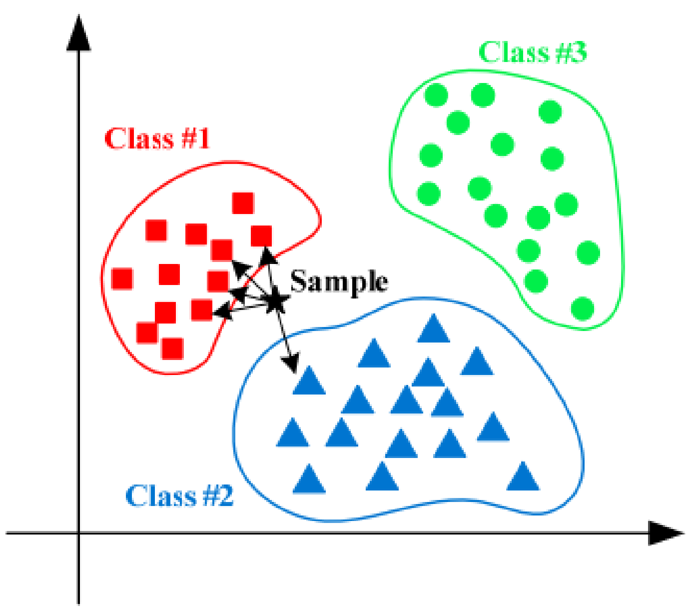


Figure 5. Description of the K-NN algorithm [2]

KNN is one of the simplest data mining classification methods. In KNN classification methods, each training sample is described as a dimensional space S according to the value of each of its features S. Then the test sample is represented in the same space and its K nearest neighbours can be obtained. The class of each K nearest neighbors is counted, and the class with the largest number of "votes" is chosen as the classifier of the test sample. K nearest neighbors are usually determined by computing the Euclidean distance between the test sample and each training sample. The Euclidean distance between the Prototype Tests and the mth Training Sample, Trainm, s is determined in the Equation:

$$D = \left[\sum_{s=1}^{s} (Test_s - Train_{m,s})^{1/2} \right]^{1/2}$$

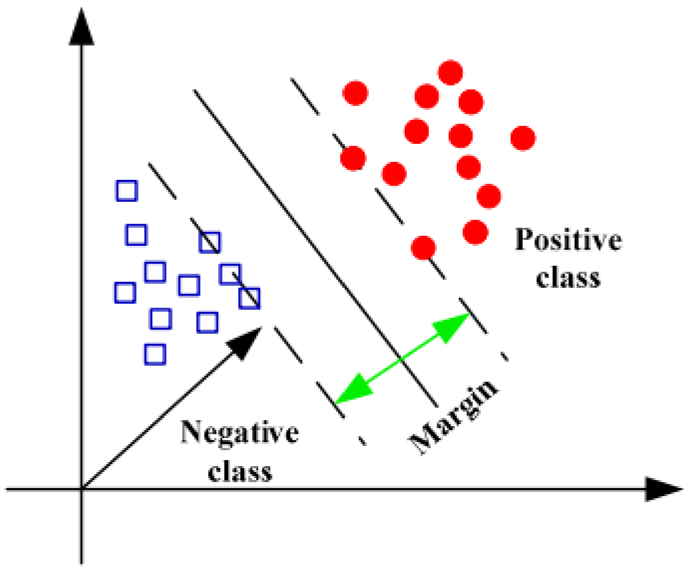


Figure 6. Description of the SVM algorithm [2]

SVM, proposed by Vapnik, is one of the most efficient and widely used classification algorithms. SVM is a machine-tilted algorithm based on statistical theory (SLT). The original idea of SVM was to use a linear decomposition hyperplane to divide the training samples into two classes. In general, two types of methods are suitable for accomplishing this criterion. The first is to find the optimal decision hyperplane that can split the two closest samples into two convex shells.

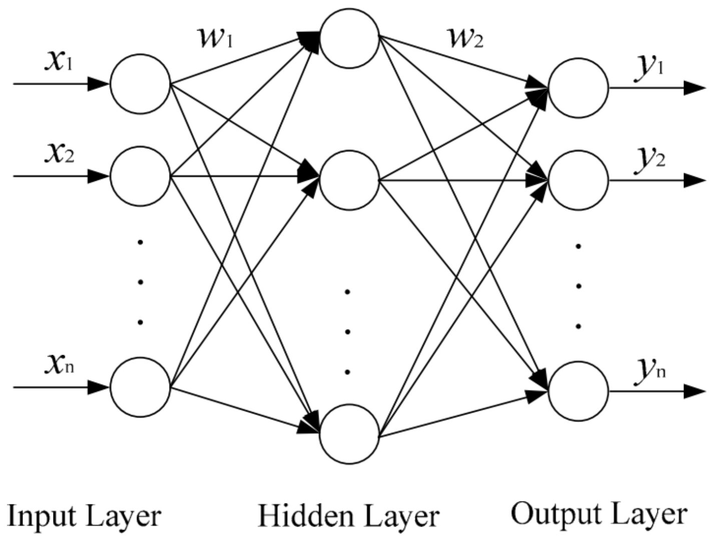


Figure 7. Description of the NN [2]

A neural network is an operating model consisting of a large number of nodes or neurons, which are interconnected. The BP neural network proposed by Rumelhart and McClelland in 1986 is a widely used neural network. The BP neural network is trained by an error propagation algorithm and is a type of multilayer feed forwarding network.

2. Methodological background

2.1 Ideas

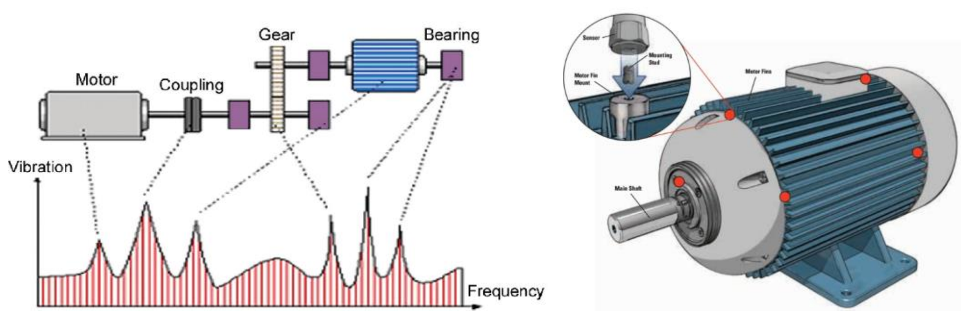
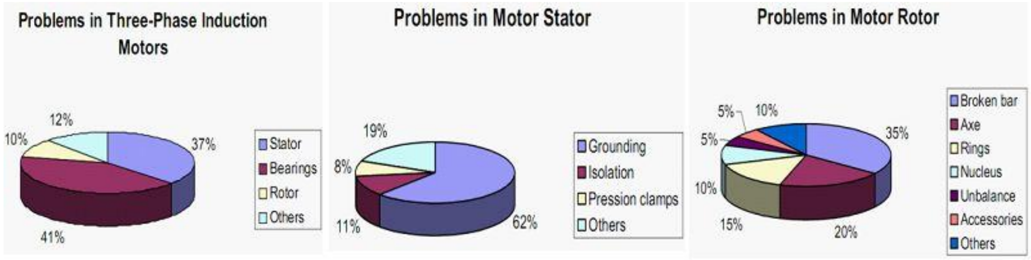


Figure 8. Vibration analysis on motor & gearbox system

We have several sensors on a bearing or motor (Figure 8): 3-axis acceleration, temperature, current. All data will be collected by controller Each signal we have a soft filter to improve the data results. The data set will be stored on the computer and the defect segmentation.

We estimate to collect 50 samples x 25 motors (good, defective: internal bearing, external bearing, Stator, Rotor...). 40% of engine types will be trained and we find a good model to fit on embedded system.

ISSN: 18158129 E-ISSN: 18151027
Volume 18, Issue 08, August, 2022



(a) (b) (c)

Figure 9. Error density graph

Since we run the model on ARM cortex M3, most of the modeling & deployment functionality will code in C Standard with no supporting framework (tensor flow, caffe...). This makes it easier to integrate on ARM Cortex embedded chips, it is independent of computer platforms.

After we complete the AI model, we will develop an IoT system, we can link many devices in the factory together. The server will monitor the devices and update the firmware via wireless or cable, it depends on them. Our board can work well with RS485/CAN/WiFi.

Since we can collect real-time and daily data, our system can see small changes in operation, and then the system can predict failure, usage time. For this case, we need to train a new (predictive) model and combine it with the old model to find the error.

2.2 Algorithm for early failure prediction (EFD)

The research team will present the data collection process, the design of an embedded board that collects and identifies and predicts errors early. The algorithm platform will be put on the ARM cortex to run independently of the computer, which is both flexible and close to the device, reducing the computer's load.

a) Data collection method

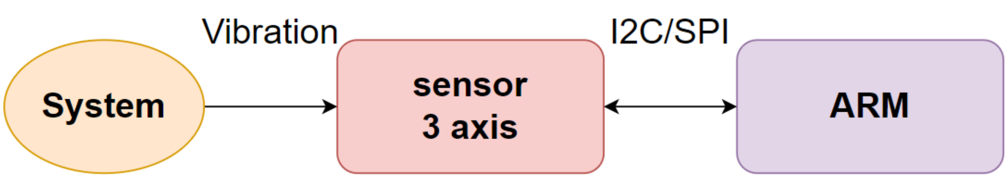


Figure 10. Data collection process

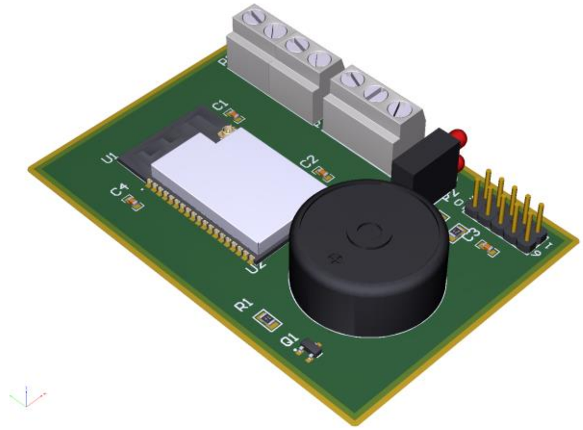


Figure 11. AIoT device collects data

Agricultural systems often do not have sensors or feedback activity values available, so it is necessary to integrate contactless sensors for better data collection dynamically.

Collected data can be ML processed on the embedded system directly without computer intervention, this helps to display visual alarms right at the device scene, simple and convenient integration.

b) Direct identification method [1], [2], [6- 9], [11]

The common faults of the motor can be classified and identified by frequency as shown in Figure 8. Thus by frequency analysis, we can easily find out the error and conduct learning to predict but widen the problem. In addition, we can also learn the vibration parameter over time to advance the geometry through the R-CNN.

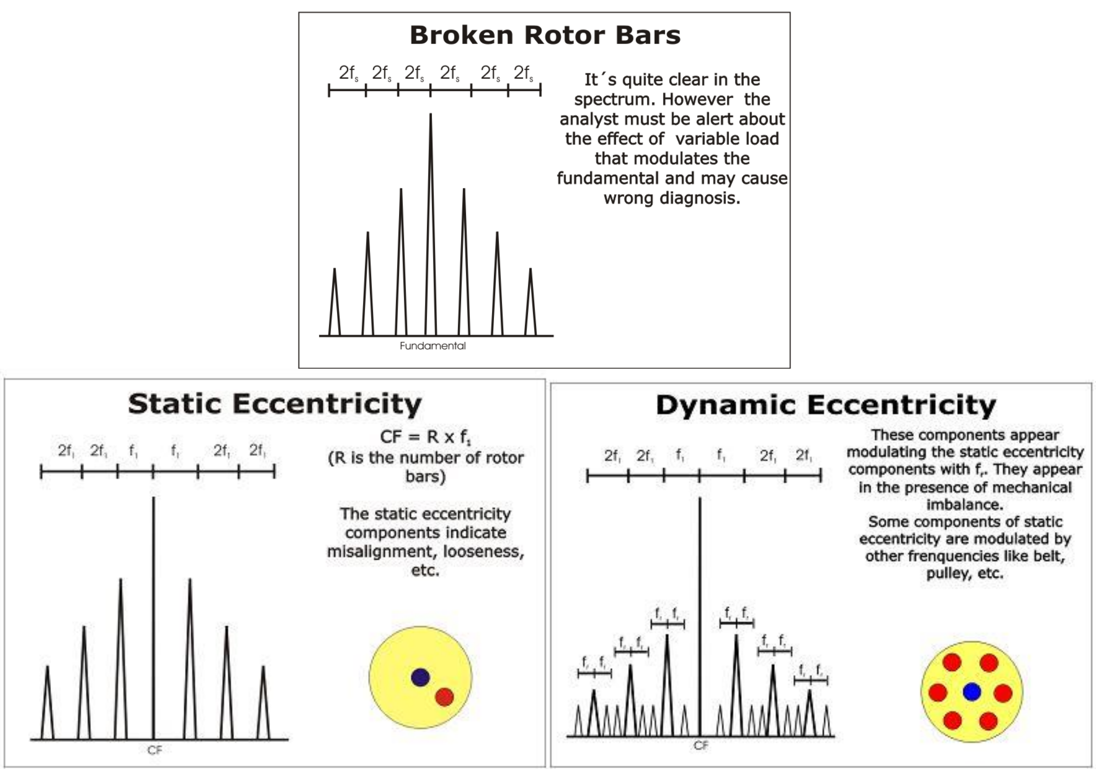


Figure 12. Fault analysis on motor

The proposed algorithm model:

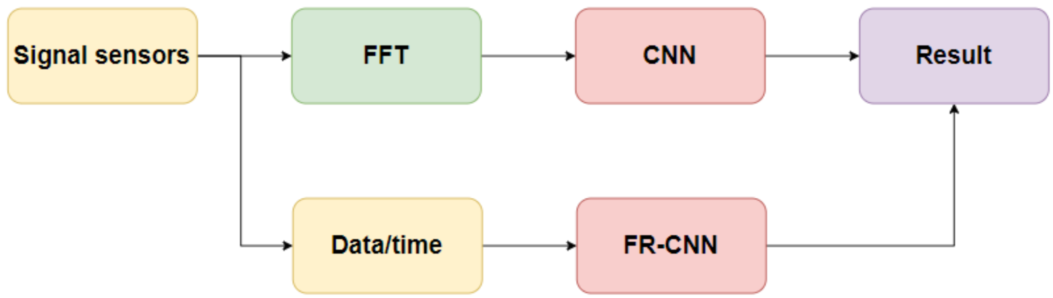


Figure 13. Block diagram of the algorithm used

The data is analyzed into two basic blocks: the temporal data block & the spectrum analysis data block.

Data over time will be processed by FR-CNN, which is the strength of time series forecasting algorithms.

FFT data we conduct CNN to highlight the difference and FC to give analysis results.

3. Conclusion

In terms of economic significance, the early detection of agricultural conveyor faults helps to significantly reduce production costs and avoid waste waiting for maintenance and repair, which current agricultural systems often encounter.

The system will manage remotely, centralize all information on the system, easily plan production, maintenance, and long-term planning issues, and avoid the risk of incidents occurring soon.

About the meaning of the algorithm: by FFT, convolutional CNN, we can put the recognizer on the embedded system to run independently and early warning of problems; this reduces the marginal cost of setting up industrial computer systems. Development direction, in the future, the group will research and apply to analyze independent machine systems, not just stop at gearboxes and motors. Since each system will have its operating frequency, the system will recognize the anomaly and make a prediction when there is a different frequency.

Acknowledgments:

This research is funded by Murata Science Foundation (MSF) under grant number 21VH03.

We acknowledge the support of time and facilities from Ho Chi Minh City University of Technology (HCMUT), VNU-HCM, for this study.

4. References

[1] https://software.intel.com/content/www/us/en/develop/topics/iot/reference-implementations/motor-defect-detector.html

[2] Wei Y, Li Y, Xu M, Huang W. A Review of Early Fault Diagnosis Approaches and Their Applications in Rotating Machinery. Entropy (Basel). 2019 Apr 17;21(4):409. doi: 10.3390/e21040409. PMID: 33267123; PMCID: PMC7514898.

[3] Lennart Ljung, System Identification- Theory for the user, ISBN-10: 0136566952

[4] Rolf Johansson, System Modeling and Identification, ISBN-10: 0134823087

[5] G. Schreck, G. Thiele, A. Fey and J. Krüger, "Robust system identification for hysteresis-controlled devices using SINDy," IECON 2020 The 46th Annual Conference of the IEEE Industrial Electronics Society, 2020, pp. 4820-4824, doi: 10.1109/IECON43393.2020.9254626.

[6] Lu C, Wang Y, Ragulskis M, Cheng Y. Fault Diagnosis for Rotating Machinery: A Method based on Image Processing. PLoS One. 2016 Oct 6;11(10):e0164111. doi: 10.1371/journal.pone.0164111. PMID: 27711246; PMCID: PMC5053415.

[7] Wu Z, Zhang Q, Wang L, Cheng L, Zhou J. Early Fault Detection Method for Rotating Machinery Based on Harmonic-Assisted Multivariate Empirical Mode Decomposition and Transfer Entropy. Entropy (Basel). 2018 Nov 13;20(11):873. doi: 10.3390/e20110873. PMID: 33266597; PMCID: PMC7512455.

[8] Chen Y, Zhang T, Zhao W, Luo Z, Lin H. Rotating Machinery Fault Diagnosis Based on Improved Multiscale Amplitude-Aware Permutation Entropy and Multiclass Relevance Vector Machine. Sensors (Basel). 2019 Oct 18;19(20):4542. doi: 10.3390/s19204542. PMID: 31635428; PMCID: PMC6832993.

[9] Huang W, Sun H, Wang W. Resonance-Based Sparse Signal Decomposition and its Application in Mechanical Fault Diagnosis: A Review. Sensors (Basel). 2017 Jun 3;17(6):1279. doi: 10.3390/s17061279. PMID: 28587198; PMCID: PMC5492044.

[10] Biggio L, Kastanis I. Prognostics and Health Management of Industrial Assets: Current Progress and Road Ahead. Front Artif Intell. 2020 Nov 9;3:578613. doi: 10.3389/frai.2020.578613. PMID: 33733218; PMCID: PMC7861342.

[11] E. L. Bonaldi, L. E. de Lacerda de Oliveira, J. G. B. d. Silva, G. Lambert-Torresm, and L. E. B. d. Silva, "Predictive Maintenance by Electrical Signature Analysis to Induction Motors", in Induction Motors - Modelling and Control. London, United Kingdom: IntechOpen, 2012 [Online]. Available: https://www.intechopen.com/chapters/40906 doi: 10.5772/48045

[12] https://docs.xilinx.com/r/en-US/xapp1356-fft-ai-engine/FFT-on-a-Single-AI-Engine