RESEARCH ARTICLE

Fault Detection in Rotating Machinery Based on Sound Signal Using Edge Machine Learning

RASHAD R. SHUBITA [ID], **AHMAD S. ALSADEH** [ID], **AND ISMAIL M. KHATER** [ID]

Department of Electrical and Computer Engineering, Faculty of Engineering and Technology, Birzeit University, Birzeit P627, Palestine

Corresponding authors: Ahmad S. Alsadeh (asadeh@birzeit.edu) and Ismail M. Khater (ikhater@birzeit.edu)

ABSTRACT Fault detection at the early stage is very important in modern industrial processes to avoid failure with life-threatening results and to reduce the cost of maintenance and machine downtime. In this paper, we present a workflow for building a fault diagnosis system based on acoustic emission (AE) using machine learning (ML) techniques. Our fault diagnosis approach is implemented on an embedded device with the internet of things (IoT) connectivity for real-time faults detection and classification in rotating machines. The achieved accuracy for our approach with a fine decision tree ML model is 96.1%.

INDEX TERMS Fault diagnosis, condition monitoring, acoustic emission, time-domain analysis, frequency domain analysis.

I. INTRODUCTION

Machines have become an integral part of all industries and applications. They are also becoming increasingly essential due to continuous technological advancements, which improve their efficiency and reliability. However, the machine will eventually fail to do its functionality for several reasons beyond one's control, such as mechanical wear and tear issues, including but not limited to bearing failure, metal fatigue, and corrosion.

Previously, sound monitoring was used by an experienced person for fault diagnosis in running machines; however, this approach depends on the experience of the operator and is not efficient [1]. Advancements in technology and signal processing algorithms tend to automate this process and make it more accurate. As a result, many condition monitoring (CM) systems for fault diagnosis have been introduced.

Fault diagnosis using artificial intelligence and machine learning techniques grasped the great attention of researchers and many research papers published in this field [2], [3], [4], [5], [6], [7], [8], [9], [10]. For instance, Tran and Lundgren [9] introduced a fault detection technique using a short-time Fourier transform (STFT) scheme. Gundewar and Kane [10] discussed the importance of sound signal-based fault classification and detection in bearings.

The associate editor coordinating the review of this manuscript and approving it for publication was Mauro Tucci [ID].

In this paper, we propose an accurate machine learning algorithm in a rotating machine, such as a drill with different configuration parameters for bearing fault, gear fault, and fan fault. Fault diagnosis based on the sound signal that can be run on industrial or consumer products using edge machine learning. We achieve accurate and reliable real-time fault detection for the intended machine. Edge machine learning (edge ML) is defined as the capability of running ML models locally to edge devices, far from the cloud or a big data center. Edge ML is mainly used when raw data is captured from a source far from the cloud or data center, which meets some requirements, such as real-time low-latency predictions, weak connectivity to the cloud, sending large data, and legal restrictions/privacy of sending data to any external cloud.

To diagnose and identify machine faults, we capture the acoustic signal of a machine using a micro-electromechanical systems (MEMS) microphone and then, analyze it in the time domain and frequency domain to extract the required features. The importance of features is ranked based on a one-way analysis of variance (ANOVA). The selected features are then used to train different machine learning techniques, such as fine decision tree, k-nearest neighbor (KNN), support vector machine (SVM), bagged trees ensemble, and naïve Bayes classifiers. The machine learning model is then deployed on an edge ML device for fault detection.

TABLE 1. Related work summary.

	Altaf et al. [5]	Kiran et al. [11]	Gundewar and Kane [10]	Liu et al. [12]]
Recorded signal	Sound	Vibration	Vibration	Sound
Sensor used	Microphone	Piezoelectric	Accelerometer	Microphone
Installation	Easy - Does not need to be mounted directly on the machine	Hard - Must be mounted on engine gearbox	Hard - Must be mounted directly on the machine gearbox	Easy - Does not need to be mounted on engine gearbox
Used techniques	K-Nearest Neighbor (KNN) Support Vector Machine (SVM) Kernel Linear Discriminant Analysis (KLDA) Sparse Discriminant Analysis (SDA)	Decision Tree Artificial Neural Network (ANN)	Artifcial neural network Discriminant classifer Support vector machine Radial basis function	Short-Time Fourier Transform (STFT) Deep Learning
Selected feature	Average FFT RMS for PSD STFT	Statistical features, such as mean, median, and Kurtosis	Statistical features such as RMS, kurtosis, standard deviation variance, etc.,	Automatically using deep learning
Monitored part	Bearing	Bearings and gears	Bearings	Bearings
Edge ML/DL considerations	No	No	No	Yes
Best achieved accuracy	**0.961**	**0.855**	**0.99**	**0.962**

II. RELATED WORK

Many methods have been proposed to automate machinery fault detection. Table 1 summarizes the closely related work.

Altaf et al. [5] presented a method for fault classification using different approaches of machine learning techniques, such as KNN, SVM, and kernel linear discriminant analysis (KLDA) for bearing fault detection and classification based on a sound signal. The importance of the research comes from applying non-traditional techniques for fault classification and detection to achieve easy and fast maintenance. While the major contribution other than applying non-traditional techniques is using sound signal instead of vibration for remote diagnosis since a vibration transducer cannot be mounted on the machine. However, the authors only targeted bearing fault in their test machine.

Kiran et al. [11] presented an engine gearbox fault diagnosis using an artificial neural network (ANN) technique for two engine part bearing and gears based on vibration signal. The achieved accuracy is 85.5% in their model. Also, a decision tree is used for feature selection. Knowing the problem in an engine gearbox, which is considered a complex machine with many moving parts inside it, makes maintenance faster and easier. The major contribution of this research is taking a completely complex machine for diagnosis, not just a simple laboratory test on a signal part. However, the combustion vibrations are not considered since the gearbox is rotated using a direct current motor, not as normally a large vibration amplitude presented when the combustion engine is on. In addition, a small bandwidth is captured just between 1 to 5kHz.

Liu et al. [12] presented an approach for fault diagnosis of bearing faults based on a sound signal with non-traditional feature extraction methods with a deep learning approach. They replaced time-consuming feature selection and complex feature extraction methods with an automatic deep learning approach for fault diagnosis. Therefore, solve the main challenge of fault diagnosis, which is observing the distinguishable fault features, which can then be forwarded for training the machine learning model. However, the proposed approach is not sufficiently satisfactory in real-time, since short-time Fourier transform (STFT) and spectrogram consume a large amount of memory footprint when doing matrix operations.
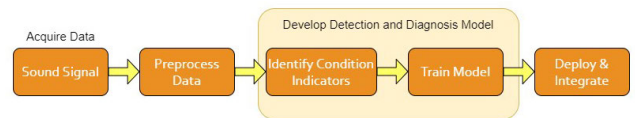

FIGURE 1. Fault detection workflow.

Also, accuracy and efficiency can be influenced by changes in working conditions.

Gundewar and Kane [10] investigated experimental bearing faults under four conditions: healthy, ball defect, outer race defect, and cage. Three different vibration signal conditions are collected as follows: filtered vibration signal, raw vibration signal, and wavelet-based denoised vibration signal. Using a neural network and discriminate classifier, Gundewar and Kane achieved very high accuracy, up to 99.58%. Although the accuracy is almost 100%, the main problem is the computational time (1.6 sec performed using MATLAB software with system configuration of intel(R) Core(TM) i7-10700CPU @2.9 GHz, 16 GB RAM), which is way more powerful than any low-cost and low-power edge device, such in our case (i.e., STM32F407 with 1-Mbyte Flash memory and 192-Kbyte RAM running @100 MHz). The high-end hardware used in Gundewar and Kane's work shows why it's cannot meet the requirement of real-time edge fault diagnosis systems.

III. METHODOLOGY

Our research is done through two phases, the first phase is the development phase, where machine sound signal is acquired using a smartphone and the reaming steps are done on a non-embedded device using MATLAB. In the second phase, the selected model is deployed and integrated into the embedded device as a sensor node. The workflow of our work is demonstrated through the following steps: *Machine Sound Signal Acquisition, Preprocessing Data, Identifying Condition Indicators, Training the Model, and Deploying & integrating the model*, as depicted in Figure 1.

A. MACHINE SOUND SIGNAL ACQUISITION

Sound is recorded at different conditions of the drill machine. We classified the drill sound signals into the following four classes: healthy, bearing fault, gears fault, and fan fault. Our

FIGURE 2. Data acquisition setup.

generated dataset of the recorded sound signals is shared at GitHub [13]. A smartphone is used as a data acquisition card to record sound signals of the drill machine at a sampling frequency of 48 kHz, according to the Nyquist-Shannon theorem, to cover all audible ranges and with oversampling by 4 kHz. The phone is placed about 10 *cm* away from the sound source as shown in Figure 2. For the experiment on the edge device and better accuracy of machine learning algorithms, a smart sensor node is used to acquire sound and store it on a secure digital (SD) card for ML algorithms training.

B. DATA PREPROCESSING

Data preprocessing is the most important step in fault diagnosis using sound signals since at this stage raw data must be noise-free as much as possible to go for further analysis. First, by applying the Hanning window, since it is the best window for unknown signals to obtain more realistic results when performing fast Fourier transform (FFT) [14]. It is essential to use the appropriate size of the window because it will affect directly the accuracy of classification [15]. We Select 2048 points as a trade-off window size to improve the accuracy of the classification and memory limitation in embedded device with 75% overlapping. Then we use a digital bandpass filter to remove the DC offset from the signal at 20 Hz and 20 kHz cut-off frequencies as an anti-aliasing filter.

C. CONDITION INDICATORS IDENTIFICATION

We analyze the drill sound signals in both time and frequency domains. For each data window, in total, 16 features are extracted to provide insight into system faults, where 6 features are extracted from the frequency domain and 10 features are extracted from the analysis of the signal in the time domain. The time domain features are listed as follows:

1) Root mean square (RMS) is defined as the square root of the arithmetic mean of the squares of the continuous signal values. See Equation 1 for the formal definition of the RMS measure [16]
2) Mean value (μ) is defined as the average of the continuous signal values. See Equation 2 for the formal definition of the Mean measure
3) Shape factor is defined as the ratio of RMS value to the mean of the absolute value of the signal
4) Kurtosis measures how outlier-prone a distribution. It is defined as in Equation 3 when faults are developing, the number of outliers increases
5) Skewness measures the asymmetry of data around the sample mean, it is defined as in Equation 4, since fault

can impact distribution symmetry, the level of skewness will increase
6) The impulse factor (IF) is used to compare the height of the signal peak to the mean level of the signal, defined as in Equation 5
7) The crest factor (CF) is defined as the peak value divided by the RMS value, since faults can be observed by changes in the peakiness of the signal before they become observable in the energy representation by RMS. CF is used for early warning of faults when they developed
8) The margin factor (MF) represents the ratio of the signal peak difference to the root amplitude
9) Variance gives a measure of the deviation of the signal from its mean value
10) The median measure is robust to outliers/noise in the signal and indicates the middlemost value of the given data, which separates the higher half of the data from the lower half

For the frequency domain sound signal features, 6 measures are extracted, which represent the values and locations of local maxima in a signal as listed below:

1) Peak1: is the largest amplitude of the extracted frequencies in the signal
2) Peak2: is the second largest amplitude of the extracted frequencies in the signal
3) Peak3: is the third largest amplitude of the extracted frequencies in the signal
4) PeakLocs1: is the frequency value of the largest amplitude
5) PeakLocs2: is the frequency value of the second largest amplitude
6) PeakLocs3: is the frequency value of the third largest amplitude

The range of frequencies contained by the preprocessed sound signal is calculated using the fast Fourier transform (FFT). After all features are extracted, a feature selection based on ANOVA is used to rank the 16 features by their importance. Peak1 is the most important feature, then variance, RMS, Peak3, Peak2, kurtosis, shape factor, and so on as shown in Figure 3.

$$RMS = \sqrt{\frac{1}{N} \sum_{i=1}^{N} x_i^2} \qquad (1)$$

$$\mu = \frac{1}{N} \sum_{i=1}^{N} x_i \qquad (2)$$

$$X_{kurt} = \frac{\frac{1}{N} \sum_{i=1}^{N} (x_i - \mu)^4}{[\frac{1}{N} \sum_{i=1}^{N} (x_i - \mu)^2]^2} \qquad (3)$$

$$X_{skew} = \frac{E(x - \mu)^3}{\sigma^3} \qquad (4)$$

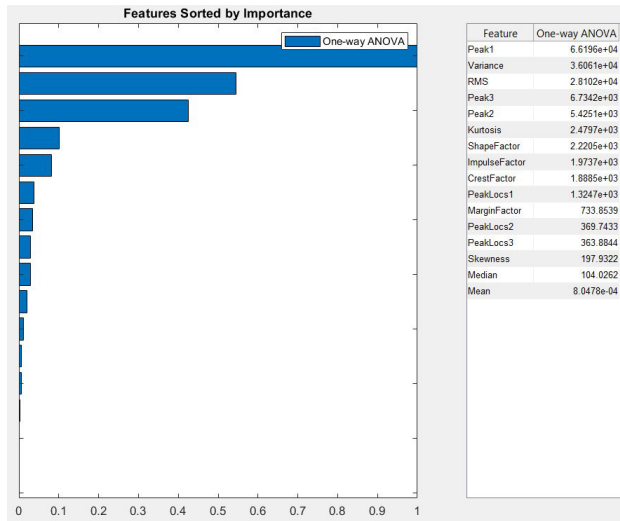$$IF = \frac{X_p}{\frac{1}{N} \sum_{i=1}^{N} |x_i|} \qquad (5)$$

FIGURE 3. Features ranking using one-way ANOVA.

where,

x: is the vector of sound signal values in the selected window size.

N: is a number of samples within window of size 2048, i.e., $N = |x|$.

σ: is the standard deviation of signal values of the vector x.

X_p: is the peak value.

Figure 4 shows a histogram for variance, RMS, and median for all the fault classes. Each fault is represented by different color as depicted in the legend of Figure 4a. Where, Fault ID 1 is off condition with white, pink, and brown noise added to the environment; Fault ID 2 is a healthy condition; Fault ID 3 is bearing fault as depicted in Figure 5a; Fault ID 4 is fan fault as depicted in Figure 5b; and Fault ID 5 is gear fault as depicted in Figure 5c. To ensure that ranking is correct, a histogram plot is used. If the features are not overlapped as in variance and RMS, this means it is a good feature that can help the model to classify the faults. For insignificant features, such as the median, the faults are overlapped in the histogram.

D. TRAIN MODEL

Selected features after the ranking step are ready to be forwarded to machine learning models. Different ML techniques are used with different configurations and a different number of features. Since the dataset is not large, we choose a trade-of value of cross-validation with $k = 5$. The best accuracy achieved is 97.4% bagged trees ensemble classifier, followed by quadratic SVM with 97.2%, fine decision tree with 96.8%, naïve Bayes with 94%, and KNN with 93.5%. A 2D scatter plot is used to observe the relationships between a pair of features, the two axes are two different features, and each observation is plotted based on these two-feature values as depicted in Figure 6.

A confusion matrix is used to evaluate the performance of the classification models, as shown in Figure 7, where the
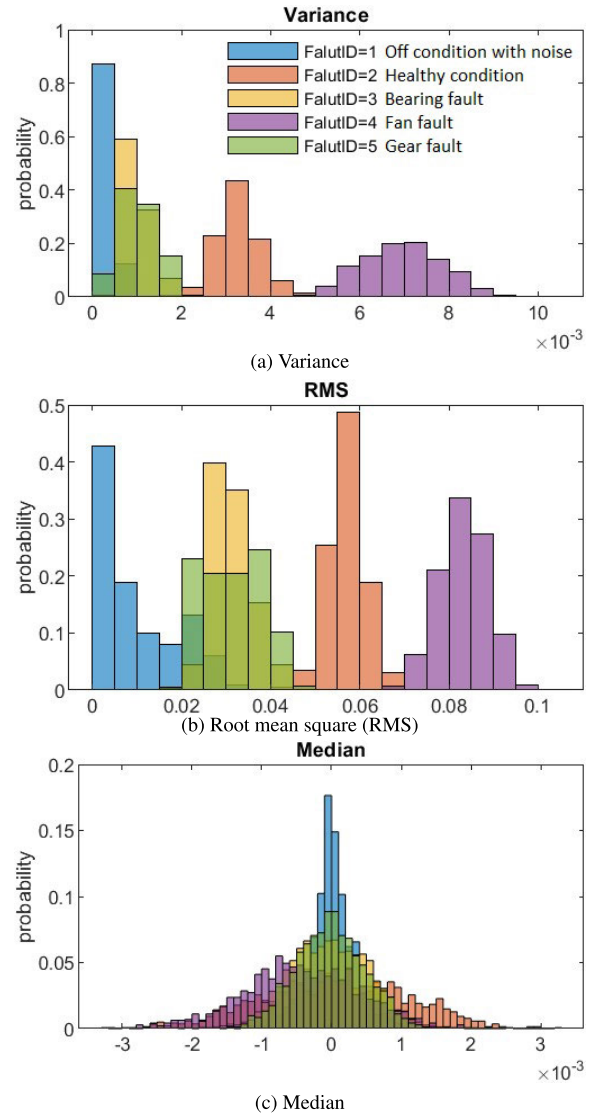

FIGURE 4. Top ranked feature distributions.

raw represents the true class (numbered as the Fault ID), and the column represents the predicted class; as a result, the diagonal observations are correctly classified samples and off-diagonal are the misclassified ones. It is noticed that when removing mean, skewness, and median features, the accuracy of the fine decision tree model does not increase. Also, most misclassification happens during the bearing fault.

IV. EXPERIMENT AND RESULTS

We select the CROWN power tool (CT10128 drill) to study the fault diagnosis of the mechanical components used in a commercial mechanical product. The details about the experimental test setup and procedure are demonstrated in the following subsections.

A. EXPERIMENT SETUP

Figure 8 shows the actual photograph of the experiment setup. The experiments were conducted on a power drill (CT10128) with extra faulty parts for fault injection into the machine.
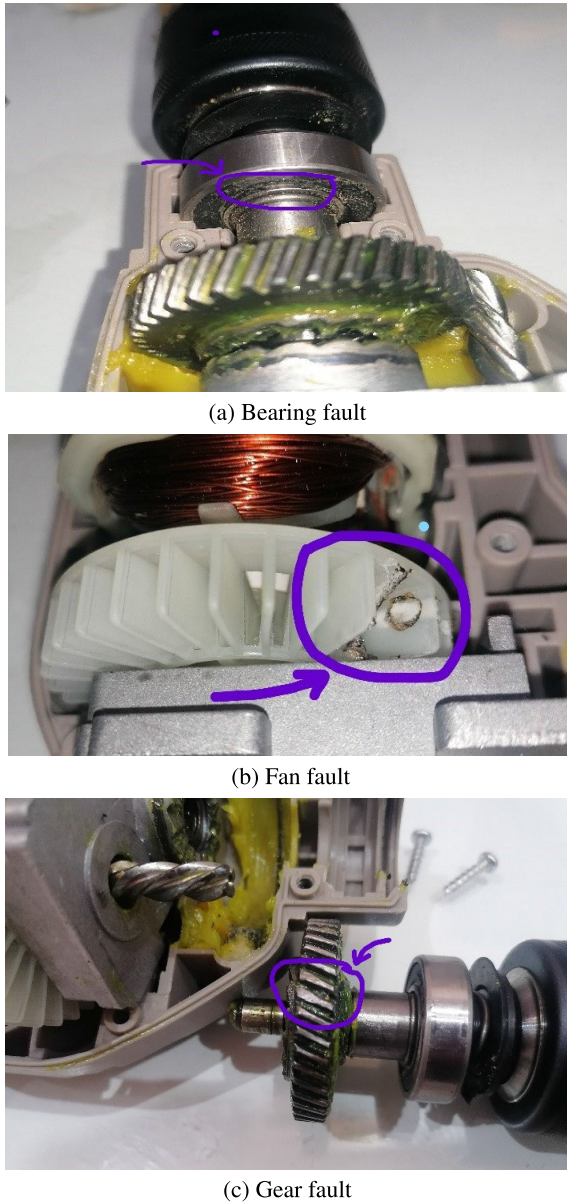
(a) Bearing fault

(b) Fan fault

(c) Gear fault

FIGURE 5. Fault cases.

(a) Scatter plot for RMS vs. Peak 1 value

(b) Scatter plot for variance vs. kurtosis

FIGURE 6. Scatter plots.

FIGURE 7. Confusion matrix for the ensemble of bagged trees classifier.

A smart sensor node is built to acquire a sound signal using a MEMS microphone (MP45DT02) connected to an ARM®Cortex®-M4 microcontroller with DSP capability (STM32F407VG). Data preprocessing and feature extraction for the selected features are performed for each data buffer (window size) having a real-time fault diagnosis node. Then, the prediction model results are then shown on the dashboard.

Using MATLAB for feature extraction, then implementing the trained model on the edge device decreases the accuracy because feature extraction performance on MATLAB is different from the edge device. Therefore, we store the extracted features using the edge device (which does not use MATLAB functions) on an SD card, then train the ML model on these features and deploy the trained ML model back to the edge device.
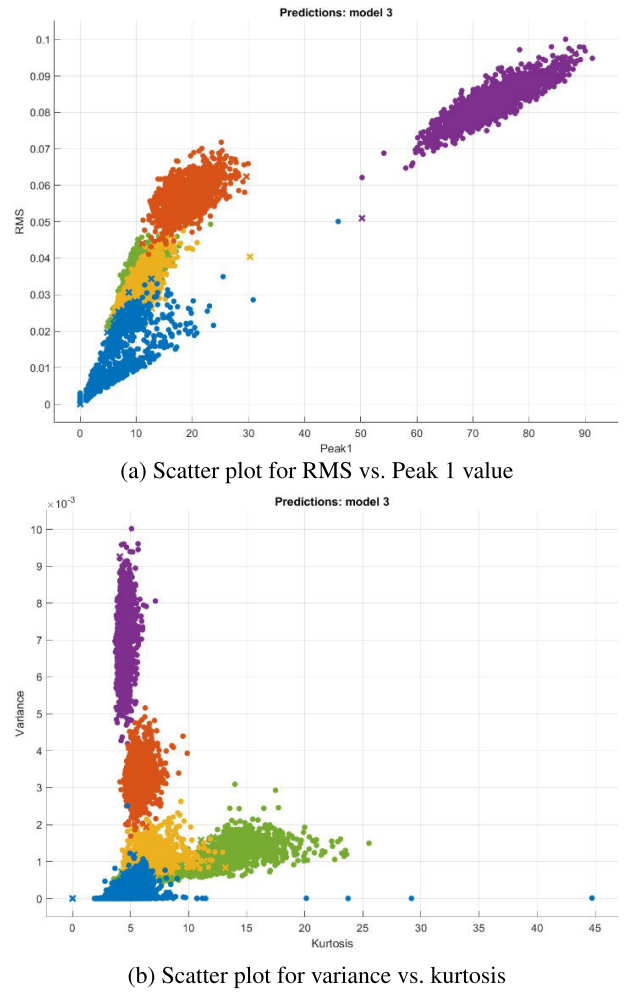
For better optimization of the hardware resources, we use a common microcontroller software interface standard to
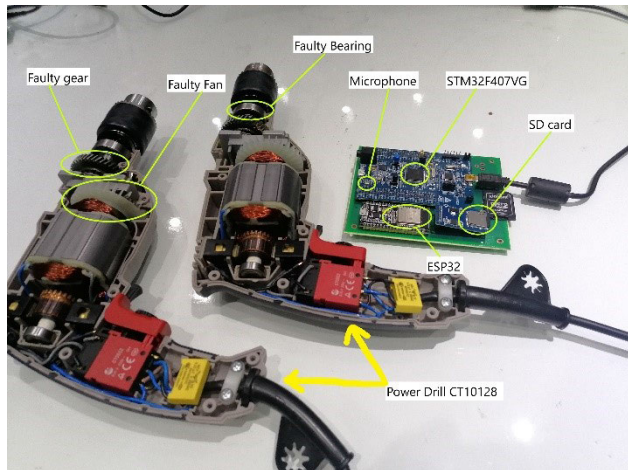
FIGURE 8. Drill machine fault diagnosis setup.

```c
483    /* calculate Max Value */
484    arm_max_f32(OutputValues,PCM_FINAL_BUFFER_LENGTH ,&MaxVal ,&Idx );
485
486    /* calculate Absolute Value */
487    arm_abs_f32(OutputValues,Abs,PCM_FINAL_BUFFER_LENGTH);
488
489    /* calculate Mean Value for Absolute Vector */
490    arm_mean_f32(Abs,PCM_FINAL_BUFFER_LENGTH,&MeanAbs);
491
492    /* calculate Variance Value */
493    arm_var_f32(OutputValues,PCM_FINAL_BUFFER_LENGTH,&SigVariance);
494
495    /* calculate RMS value */
496    arm_rms_f32(OutputValues,PCM_FINAL_BUFFER_LENGTH,&SigRMS);
497
498    /* calculate Crest Factor value */
499    SigCrestFactor = MaxVal/SigRMS;
500
501    /* calculate Shape Factor value */
502    SigShapeFactor = SigRMS/MeanAbs;
503
504    /* calculate Impulse Factor value */
505    SigImpulseFactor = MaxVal/MeanAbs;
506
507    /* calculate Margin Factor value */
508    SigMarginFactor = MaxVal/(MeanAbs * MeanAbs);
```

FIGURE 9. DSP functions for features extraction in time domain.

develop a real-time digital signal processing (CMSIS-DSP)[1] used mainly for data preprocessing and feature extraction as shown in Figure 9. MATLAB Coder is used to generate a *C* function for features extraction in the frequency domain for a given FFT spectrum with pre-defined prominence value for peak extraction and a *C* function for the selected machine learning algorithm. Morover, for extra optimization in memory footprint for frequency domain feature extraction in MATLAB function, it is noticed that most peak values located (highest three) in a range 20 to 3K Hz using probability distribution methods, such as probability density function (PDF) and normal distribution fit as shown in Figure 10.

B. EXPERIMENT PROCEDURE

In this experiment, We injected artificial faults into the Drill (CT10128) components. In the case of the bearing fault, a small particle is injected inside the bearing to get almost a real fault. In the case of the fan fault, two blades are broken. While in the case of gears fault, two teeth destructive pitting failures are introduced.
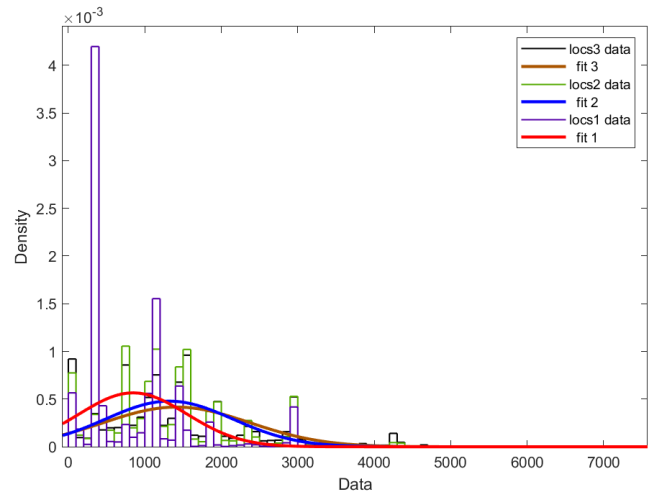
FIGURE 10. Distribution of data in FFT spectrum for all peaks.

The smart node has two modes, the 'run mode' and 'recording mode'. Five states are recorded in recording mode and their associated features are logged to an SD card, then the node switch to the run mode for testing. The algorithm and testing results are discussed in the next subsection IV-C.

C. FAULT DETECTION ALGORITHM

Figure 11 shows the complete workflow in the implementation phase. First, the record mode is selected to collect and log features to an SD card, each record is one minute long and each record represents a fault state. As a result, more than three faults can be added, such as a combined fault that may appear in the normal lifecycle of the machine. After that, use MATLAB to train the model on these collected features. The selected model is the fine decision tree model. Even the accuracy is not the best as shown in the results from the development phase, but compared to other ML techniques it runs faster and has a small memory footprint. Then update the model on the microcontroller and restart the device to select the run mode. In the run mode, the same preprocessing and feature extraction algorithms in record mode are used. Ping-pong buffering techniques are used as shown in Figure 12, to process and acquire data in parallel to achieve real-time prediction without any data loss.

The extracted features are fed to the machine learning fine decision tree model to make a prediction. To compensate for the reduced accuracy in the fine decision tree model, a simple empirical cumulative distribution function is used on a block of prediction results and gets the most repetitive prediction as a final prediction result. Then the final prediction is sent to ESP32 to send it to a dashboard on the internet.

D. EXPERIMENT RESULTS

A total of 99610 feature values are taken as input to fine decision tree ML model ($996110 = 7115$ observation $\times 14$ features for each observation, where unimportant features, such as median and mean as shown in Figure 3 removed to reduce the dimensions of features). It is noticed that the model
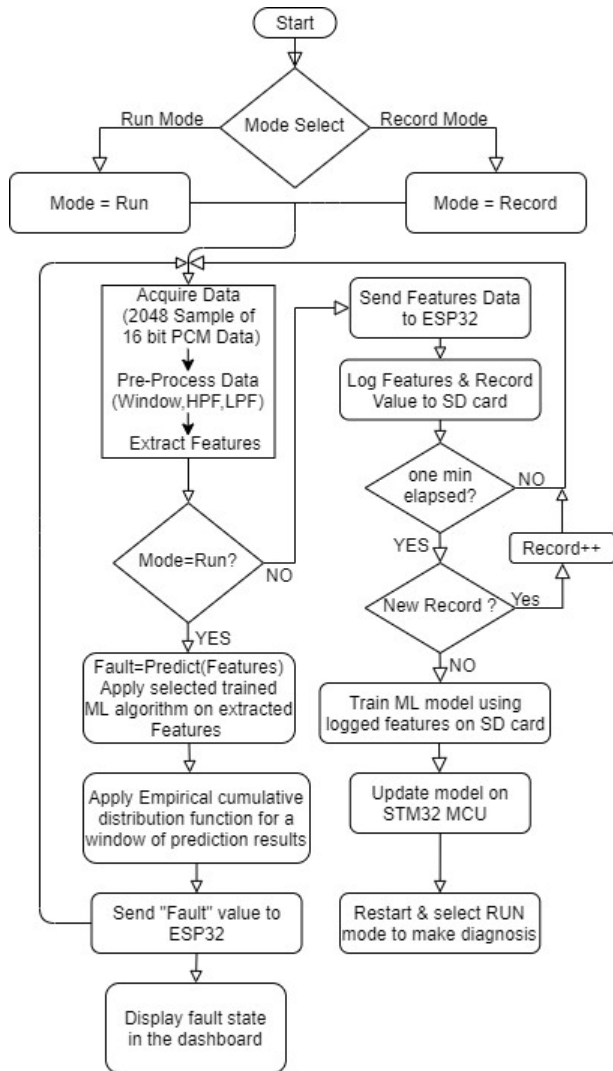
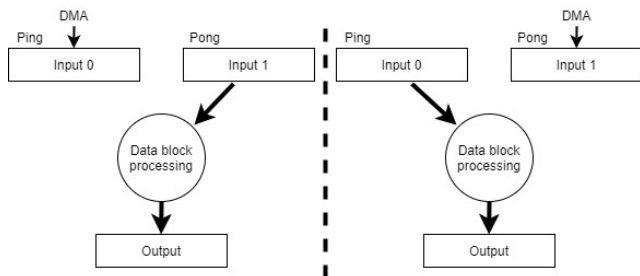**FIGURE 11. Flowchart for the implementation workflow.**

FIGURE 12. Ping-Pong Algorithm.

accuracy drops slightly from 96.8% in the development phase to 96.1% in the implementation phase as Figure 13 illustrates the confusion matrix of the fine decision tree model where most of the misclassifications happened with bearing fault. 0.7% drop in accuracy as a result of changing hardware used in both development and implementation phases and performance difference during feature extraction in hardware and MATLAB environment. However, reducing feature dimensions do not affect the accuracy significantly of the model,

FIGURE 13. Confusion matrix for a fine decision tree model.

but improves the performance in terms of classification speed. Also, the empirical cumulative distribution function improves the stability (precision) of the model.

V. DISCUSSION

Our proposed method and analysis achieve a high accuracy of 96.1%, the same as in Altaf et al. work [5], but for the different types of faults, not only for bearing fault. In Kiran et al. [11], the model ran on a gearbox as a complete machine and achieved 85.5% accuracy for gear and bearing faults without edge ML consideration and neglecting the vibration from the combustion engine. Compared to Gundewar and Kane's research [10], from the Table 1, they achieved the highest accuracy among other bearing fault diagnosis methods. Still, the main drawback of using a neural network with a discriminate classifier, is the high computational time, even running on a powerful computing device compared to an embedded device with limited resources.

As in Liu et al. research [12], deep learning can eliminate the feature selection process and achieve reasonable accuracy. But as the authors state, it consumes a large memory footprint and unstable accuracy when working conditions changes. In contrast, we use an efficient feature selection method using one-way ANOVA to enhance the ability to improve and optimize the model. We use the minimum amount of data and extract the relevant features from the row data to achieve stable accuracy and real-time performance with a simple model for edge ML.

Even though the empirical cumulative distribution function is used to improve the stability of the model, it still can be operated in real engineering applications and meet industrial requirements. A more powerful controller can meet this requirement by running the bagged trees ensemble model with 97.4% accuracy. Furthermore, before feature extraction at the preprocessing stage, many new methods, such as in Zijian et al. research [17], can enhance useful data embedded in the noise signal to improve the model's overall accuracy. This can be done in future research.

VI. CONCLUSION

In this paper, we conducted a case study for fault diagnosis of three rotating elements of commercial drill tool CT10128 (bearing, fan, and gear). Starting from the concept to the development phase where different ML approaches are used with different configurations to select the best classifier to the implementation phase using a tree classifier on an embedded device (STM32F407), where all diagnosis steps, including data acquisition, pre-processing, features extraction, and fault classification is done at edge device without the use of the cloud to enhance the security, performance of the diagnosis system, and minimize network bandwidth for a cost-effective solution [18].

For a successful edge ML fault diagnosis system, the following considerations need to be met: (i) a good amount of observation for each condition, (ii) good preprocessing algorithms, such as a window with overlapping and digital low/high pass filters, (iii) appropriate features extraction algorithm and implementation to achieve high accuracy diagnosis system. Therefore, according to our experiment result, edge machine learning can be used for diagnosing faults on a machine that has a rotating component based on a sound signal using appropriate development and implementation techniques. Furthermore, based on the accuracy and stability of the proposed method, it can be recommended for practical applications for online machine condition monitoring.

ACKNOWLEDGMENT

REFERENCES

[1] R. A. Collacott, *Mechanical Fault Diagnosis and Condition Monitoring*. Berlin, Germany: Springer, 2012.

[2] T. Holroyd and N. Randall, "Use of acoustic emission for machine condition monitoring," *Brit. J. Non-Destructive Test.*, vol. 35, no. 2, pp. 75–78, 1993.

[3] D. Mba, A. Cooke, D. Roby, and G. Hewitt, "Opportunities offered by acoustic emission for shaft-seal rubbing in power generation turbines: A case study," in *Proc. Int. Conf. Condition Monitoring*, Oxford, U.K., 2003, pp. 2–4.

[4] Y. Kim, A. Tan, J. Mathew, and B. Yang, "Experimental study on incipient fault detection of low speed rolling element bearings: Time domain statistical parameters," in *Proc. 12th Asia–Pacific Vibrat. Conf.* Japan: Hokkaido University, 2007, pp. 6–9.

[5] M. Altaf, M. Uzair, M. Naeem, A. Ahmad, S. Badshah, J. A. Shah, and A. Anjum, "Automatic and efficient fault detection in rotating machinery using sound signals," *Acoust. Aust.*, vol. 47, no. 2, pp. 125–139, Aug. 2019.

[6] E. Kim, C. A. Tan, J. Mathew, V. Kosse, and B.-S. Yang, "A comparative study on the application of acoustic emission technique and acceleration measurements for low speed condition monitoring," in *Proc. 12th Asia Pacific Vibrat. Conf.*, 2007, pp. 1–11.

[7] D. Mba, "The detection of shaft-seal rubbing in large-scale turbines using acoustic emission," in *Proc. 14th Int. Congr. Condition Monitor. Diagnostic Eng. Manage. (COMADEM)*, Manchester, U.K., Sep. 2001, pp. 21–28.

[8] C. K. Tan and D. Mba, "Limitation of acoustic emission for identifying seeded defects in gearboxes," *J. Nondestruct. Eval.*, vol. 24, no. 1, pp. 11–28, Mar. 2005.

[9] T. Tran and J. Lundgren, "Drill fault diagnosis based on the scalogram and MEL spectrogram of sound signals using artificial intelligence," *IEEE Access*, vol. 8, pp. 203655–203666, 2020.

[10] S. K. Gundewar and P. V. Kane, "Rolling element bearing fault diagnosis using supervised learning methods- artificial neural network and discriminant classifier," *Int. J. Syst. Assurance Eng. Manage.*, vol. 13, no. 6, pp. 2876–2894, Dec. 2022.

[11] K. Vernekar, H. Kumar, and G. K. V., "Engine gearbox fault diagnosis using machine learning approach," *J. Quality Maintenance Eng.*, vol. 24, no. 3, pp. 345–357, Aug. 2018.

[12] H. Liu, L. Li, and J. Ma, "Rolling bearing fault diagnosis based on STFT-deep learning and sound signals," *Shock Vibrat.*, vol. 2016, Sep. 2016, Art. no. 6127479.

[13] R. Shubita. (2022). *Fault-Detection-Using-TinyML*. [Online]. Available: https://github.com/RashadShubita/Fault-Detection-using-TinyML

[14] National Instruments. (Dec. 2021). *Understanding FFTS and Windowing*. [Online]. Available: https://www.ni.com/en-us/innovations/white-papers/06/understanding-ffts-and-windowing.html

[15] S. M. Tayyab, E. Asghar, P. Pennacchi, and S. Chatterton, "Intelligent fault diagnosis of rotating machine elements using machine learning through optimal features extraction and selection," *Proc. Manuf.*, vol. 51, pp. 266–273, 2020.

[16] *IEEE Standard for Transitions, Pulses, and Related Waveforms*, IEEE-SA Standards Board, Standard IEEE 181-2011, 2011, pp. 1–71.

[17] Z. Qiao and X. Shu, "Coupled neurons with multi-objective optimization benefit incipient fault identification of machinery," *Chaos, Solitons Fractals*, vol. 145, Apr. 2021, Art. no. 110813.

[18] M. Nadeski. (Mar. 2019). *Bringing Machine Learning to Embedded Systems*. [Online]. Available: https://www.ti.com.cn/cn/lit/wp/sway020a/sway020a.pdf

RASHAD R. SHUBITA received the bachelor's degree in mechatronics engineering, in 2021. He is currently pursuing the master's degree in computer engineering with Birzeit University. His main research interests include machine condition-based monitoring, signal processing, embedded systems, fault diagnosis, and TinyML.

AHMAD S. ALSADEH received the bachelor's degree in electrical engineering with emphasis on telecommunication, in 2002, the master's degree in scientific computing from Birzeit University, in 2007, and the D.Eng. degree in network security from the Hasso Plattner Institute, University of Potsdam, Germany, in October 2013. He is currently an Assistant Professor with the Electrical and Computer Engineering Department, Birzeit University, where he mainly teaches IT security courses for master's and bachelor's students. His research interests include network security, applied cryptography, and systems security.

ISMAIL M. KHATER is currently an Assistant Professor with the Department of Electrical and Computer Engineering, Birzeit University. He has numerous publications within the specialty and published in reputed national and international peer-reviewed journals. He gains recognition among the honorable subject experts with the contributions made. His major research interests include utilizing computational intelligence and neural networks for wide applications in medicine, biology, physics, engineering, and many other research fields.

• • •