

Article

Application of Convolutional Neural Network to Defect Diagnosis of Drill Bits

Yongchao Yu ^{1,2} , Qi Liu ^{1,2} , Boon Siew Han ¹ and Wei Zhou ^{1,2,*} ¹ Schaeffler Hub for Advanced Research, 61 Nanyang Dr, ABN-B1b-11, Singapore 637460, Singapore² School of Mechanical and Aerospace Engineering, Nanyang Technological University, 50 Nanyang Avenue, Singapore 639798, Singapore

* Correspondence: mwzhou@ntu.edu.sg or wzhou@cantab.net

Abstract: Drilling, one of the most used machining processes, has wide application in different industrial fields. Monitoring the system health and operation status of the drilling process is essential for maintaining production efficiency. In this study, a convolutional neural network (CNN), a deep-learning method, is applied to the defect diagnosis of drill bits. Four drill bits with different health conditions were used to drill holes in an aluminum block, and a vibration sensor collected the signals. Vibration spectrograms generated using short-time Fourier transform were applied to a 2D CNN algorithm, and they were then reconstructed into a 1D data set and applied to a 1D CNN algorithm. The input data size was reduced significantly compared to the raw vibration data after the data-reconstruction process. As a result, the 2D CNN process shows a diagnostic accuracy of 97.33%. On the other hand, the 1D CNN provides a diagnostic accuracy of 96.6%, but it only requires 2/3 of the computational time required by the 2D CNN.

Keywords: defect diagnosis; convolutional neural network; condition monitoring



Citation: Yu, Y.; Liu, Q.; Han, B.S.; Zhou, W. Application of Convolutional Neural Network to Defect Diagnosis of Drill Bits. *Appl. Sci.* **2022**, *12*, 10799. <https://doi.org/10.3390/app122110799>

Academic Editor: Cem Selcuk

Received: 6 October 2022

Accepted: 21 October 2022

Published: 25 October 2022

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



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

1.1. Machine Fault Diagnosis

In today's highly automated industry, keeping track of machine condition is crucial to maintain productivity and product yield. During the operation process, any abnormal conditions will not only reduce machining accuracy and quality but also increase production costs due to tool replacement and rework. Therefore, machine condition monitoring has gained significant enthusiasm [1].

Fault diagnosis is an important part of machine condition monitoring. It affects a project's efficiency and greatly concerns the protection of the machine itself. Therefore, a great deal of effort has been expended to develop mechanical fault diagnostic methods. As artificial intelligence (AI) develops, various AI methods have been applied to machine fault diagnosis in recent years. Dai et al. developed a machine vision system combined with an image processing algorithm to monitor the progressive wear of tools and to reduce downtime [2]. The effectiveness of the prototype system and the developed algorithms were verified. Bhuiyan et al. applied an acoustic emission sensor to measure the vibration frequency in metal machining for tool condition monitoring [3].

To increase diagnostic accuracy, feature extraction is an essential process for most machine learning algorithms. It affects the diagnostic accuracy directly. Therefore, various feature extraction methods have been developed and verified. Asr et al. designed a feature extraction method based on empirical mode decomposition (EMD) [4]. The extracted features were input into a non-naive Bayesian classifier for fault diagnosis. Georgoulas et al. proposed a symbolic aggregate approximation framework and extracted features from bearing signals, and a nearest neighbor classifier was used for classification [5]. Xiong et al. demonstrated the use of a multi-fractal detrend fluctuation analysis to extract features

for bearing fault diagnosis [6]. Zhang et al. demonstrated the use of variational modal decomposition (VMD) for feature extraction in mechanical fault diagnosis [7].

Many of the methods mentioned above have achieved decent classification and diagnostic accuracy, but the feature extraction process highly depends on the experimenters' experience and the original data. The same feature extraction process and classification method may not be able to provide an accurate result when using different workloads and setups. To address this problem, Zhang et al. conducted comparative experiments in their recent research [8]. In their results, it shows that convolutional neural network (CNN) has higher adaptability than support vector machine (SVM), multi-layer perceptron (MLP), and deep belief network (DBN). In recent years, among different deep-learning algorithms, CNN has gained massive interest due to its wide range of practicality and high accuracy [9,10].

1.2. Convolutional Neural Network

CNN is a multistage neural network with multiple filtering and classification stages. The network's performance is highly dependent on the total number of layers, the number of convolution kernels, and the number of neurons in the network. The construction of a CNN model can be performed by cross stacking the convolutional layer, the activation layer, the pooling layer, and the fully connected layer [11].

CNN algorithms have been developed and applied to various fault detection applications, especially in rotation systems such as bearings and gears. Hoang et al. demonstrated a simple two-layer convolutional model for rolling bearing fault classification by converting vibration signals into gray-scale images [12]. In this study, the vibration signals were normalized into the range of $[-1, 1]$. Then, the normalized data were converted to gray-scale intensity of each pixel in the corresponding image. Using the same approach, Wang et al. converted multi-sensor vibration signals into a single image file [11]. Instead of using a gray-scale image, Wang et al. transformed the signals from three channels into a color image. Each channel corresponded to one of three colors: red, blue, and green (RGB). They then combined the three colors to produce RGB-colored images. The features of the color image enlarged the differences between the different fault signals and increased the fault detection accuracy.

Instead of directly using vibration signals, applying a signal processing method also increases the diagnostic accuracy. Janssens et al. demonstrated a 2D CNN process for bearing condition classification by applying Fourier transform to acceleration signals [13]. Ding et al. presented a novel energy-fluctuated multiscale feature mining approach based on a wavelet packet energy (WPE) image to perform a 2D CNN for spindle-bearing fault diagnosis [14]. The WPE transfer separated the signals into a series of frequency subspaces without losing the energy of the frequency components and greatly enlarged the vibration frequency characteristics. Furthermore, Verstraete et al. compared the effectiveness of several time–frequency analytical methods, including the short-time Fourier transform, the wavelet transform, and the Hilbert–Huang transform [15].

Two-dimensional CNN, as a deep-learning algorithm, shows promising diagnostic accuracy. However, 2D CNN was initially designed for image analysis; it shows some limitations in processing vibration signals because vibration signals are usually presented as time-domain 1D data. To overcome this problem, 1D CNN has been applied to machine failure detection.

One-dimensional CNN was first demonstrated in 2015 to analyze electrocardiograms and to detect human activity signals. Meanwhile, 1D CNN was also applied to bearing/gear fault detection. Using vibration signals is the most common approach to performing 1D CNN in bearing fault detection. Eren et al. used raw vibration signals as the input data to train a 1D CNN for bearing fault detection [16]. Studies also showed that real-time and rapid failure detection could be achieved by monitoring vibration signals and processes in 1D CNN [16,17].

Besides monitoring vibration signals, researchers are also looking to use other mechanical signals for 1D CNN-based classification. For example, Ince et al. used the motor's current data to train a 1D CNN for real-time condition monitoring [18]. Similarly, Khan et al. used the motor's current data as the input and developed an analytical model for inter-turn fault diagnosis by combining the 1D CNN and long short-term memory network [19]. Both 1D and 2D CNNs have shown high accuracy in classification and diagnosis. CNN retains more attributes from the original data. Therefore, it can reduce error or information loss during the feature extraction process.

In this study, a CNN-based drill bit defect diagnostic application is demonstrated. Drilling is one of the most used machining processes. During the drilling process, the abrasion of the drill bits is much greater than that of the bearings and gears. The health condition of the drill directly affects the success or failure of drilling. Additionally, drilling is one of the significant processes of rock mechanics and mining engineering. Mining engineers are interested in identifying geological information by monitoring the vibration signals from the drilling process.

In recent years, Yari et al. reported the logical relations between rock characterization and acoustic frequencies by monitoring the acoustic signals generated during the drilling operation [20]. Furthermore, by analyzing acoustic signals, various physical characteristics of sedimentary rocks and carbonate rocks can be determined [20,21]. Khoshouei et al. used the acoustic signal's root mean square (RMS) to identify the uniaxial compressive strength; the Brazilian tensile strength; the Schmidt rebound number; and the longitudinal wave velocities of sedimentary, metamorphic, and igneous rocks [22]. These approaches have significant potential for their future use in the industry for field identification and classification of rock.

However, the premise of all of these applications is to use a non-damaging drill bit. Therefore, establishing a highly effective condition monitoring and failure detection system is essential and helps to reduce unnecessary material and time consumption and protects the machine from damage. This study demonstrates a drill bit failure diagnostic system by monitoring vibration patterns during the drilling process. Vibration signals were collected using a SmartCheck system, and drill bits with different health conditions were used. Short-time Fourier transform was performed on the raw data to create a spectrogram image for each measurement data set. After processing the data, CNN algorithms were applied for defect classification and diagnosis (Figure 1). Additionally, 1D CNN and 2D CNN were used for data processing. Finally, the characteristics of the two algorithms were compared.

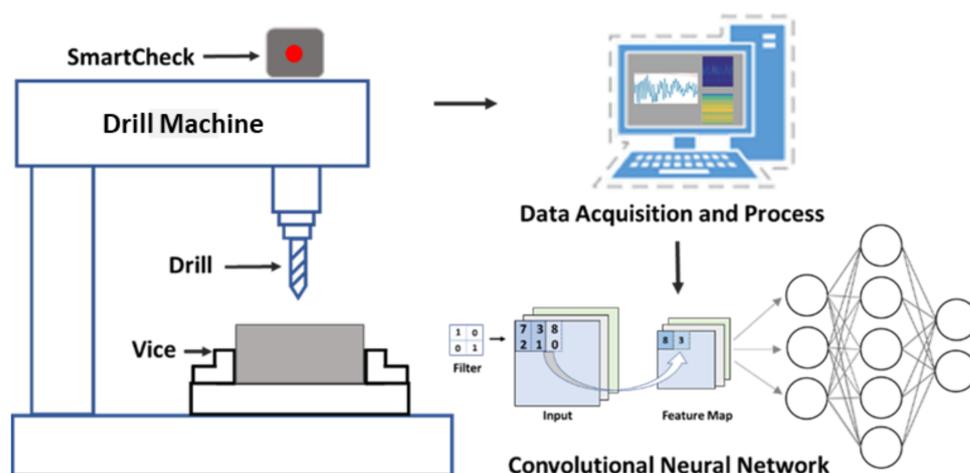


Figure 1. Schematic representation of the experimental process.

2. Experiment Setup

In this experiment, a SmartCheck system, which is a novel condition monitoring tool produced by the Schaeffler Group, was used for vibration data collection. A piezoelec-

tric accelerometer and a temperature sensor were integrated into the SmartCheck system. Figure 2a shows the image of the SmartCheck system used in this experiment. The acceleration sensor in the SmartCheck system has a ± 50 g measurement range and a frequency range of 0.8 Hz–10 kHz. This frequency range is suitable for measuring most of the machine vibrations. The sampling frequency of the SmartCheck system was set to 25.8 K for high-resolution measurement.

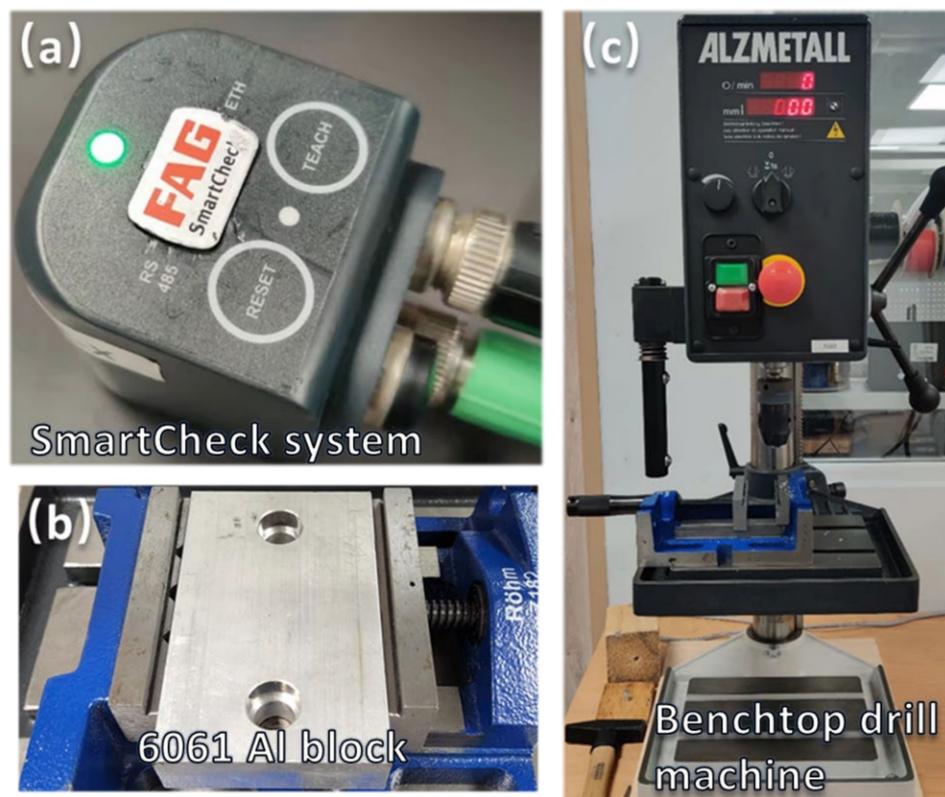


Figure 2. (a) An image of the SmartCheck system, (b) an image of the aluminum block, and (c) an image of the drill machine used in the experiment.

The sensor was attached to the top of the drill machine to measure vibrations (Figure 1). The sensor position was selected because it was the closest position to the motor of the drill machine and helped in the collection of vibration signals. Additionally, this position would not be disturbed when changing drill bits or moving samples. It ensured that the sensor was placed in the same location for all measurements and reduced measurement error.

During the measurement, the drilling speed was controlled by a control panel and was set to 500 rpm. The feeding speed was approximately 3 mm/min. Each vibration signal was continuously collected for 1.27 s, and the collecting interval of two signals was 20–25 s. Each data set contains 325,120 data points. Figure 2b presents the 6061 aluminum block used in the experiment. The aluminum was fixed on the stage by a clamp for safety purposes and to avoid displacement during the drilling process. Figure 2c shows the drill machine used in this experiment.

Figure 3 presents three drill bits with three artificially created defects and a drill bit without defects. Yellow circles show the defect positions. Figure 3a presents a notched blade edge. There is evident damage on the edge of the blade. Figure 3b presents a blunt edge in a drill bit, as shown in the yellow circle. The edge of the drill bit was polished by file tools, which made the blade not as sharp as a healthy drill bit. Figure 3c presents a defective drill bit with a blunt head. The top of the drill bit was flattened. A yellow circle highlights this in the image. These are the most common defects in drill bits caused by daily operations [23]. Lastly, Figure 3d presents a healthy drill bit without defects.

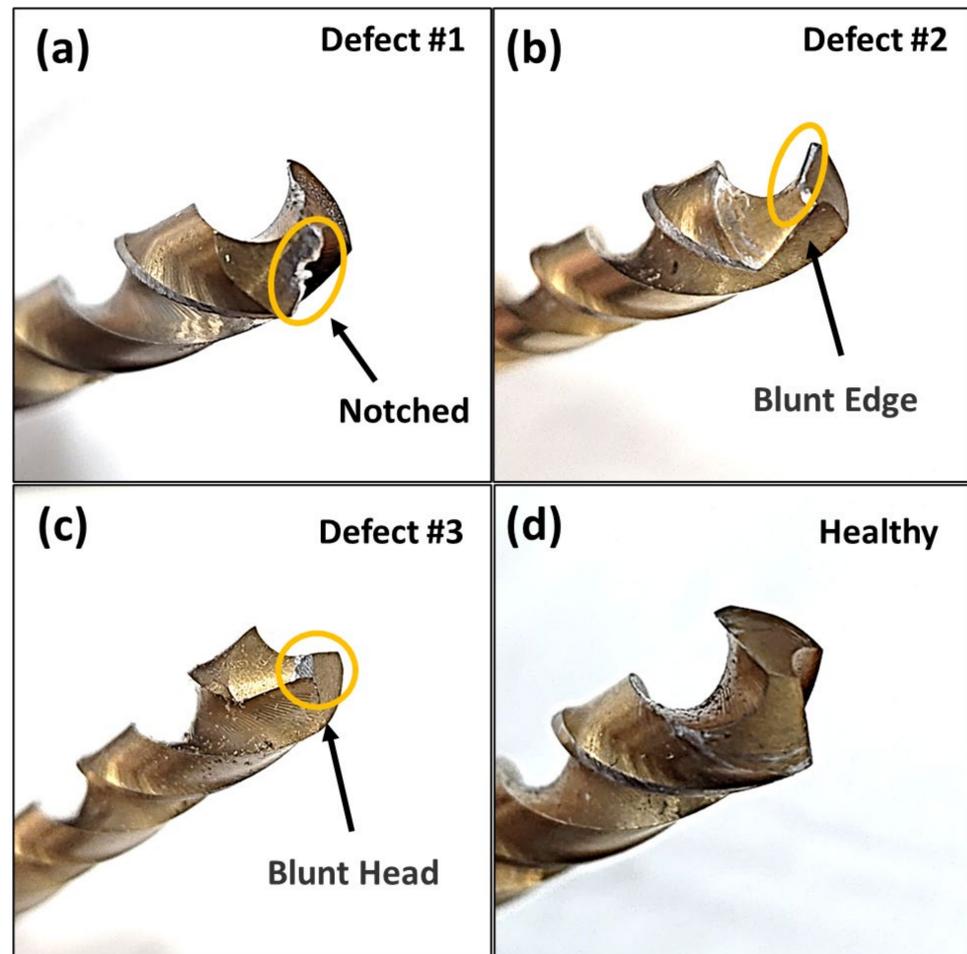


Figure 3. (a) An image of a notched drill bit, (b) an image of a blunt-edged drill bit, (c) an image of a blunt-headed drill bit, and (d) an image of a healthy drill bit without any defects.

3. Data Processing and Feature Extraction

Figure 4 presents the time-domain vibration signals of the four drill bits with different health conditions. Defects #1, #2, and #3 represent the “notched” drill bit, the drill bit with a “blunt edge”, and the drill bit with a “blunt head”, respectively. “Healthy” represents the drill bit with no damage. The figure shows the accelerometer read-out. The amplitude corresponds to the unit in g ($1 g = 9.81 m/s^2$). A larger amplitude represents a larger vibration amplitude. Each data set contains 32,770 data points. From the figure, it can be observed that the vibration signals contain a large amount of noise. It is challenging to distinguish the differences between each condition. In particular, the signal amplitude and vibration features of defect #1 and the “healthy” sample are very similar. Therefore, further data processing is needed for defect diagnosis. Although some researchers use raw vibration data to perform 1D CNN for bearing fault detection [16], it requires a large amount of computation. This study applied a short-time Fourier transform to reduce the data size and the noise effect.

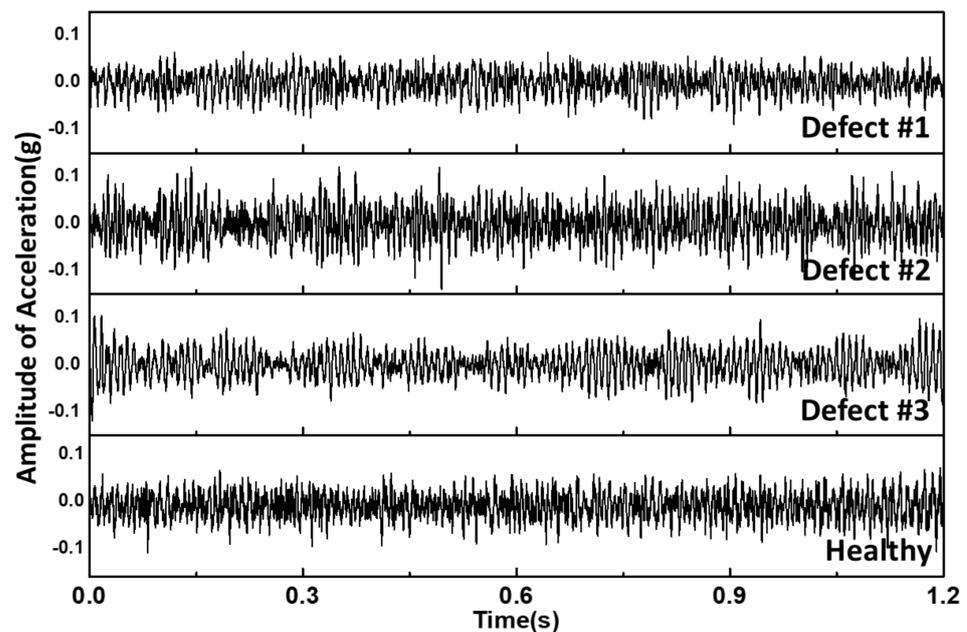


Figure 4. Read-out of vibration signals collected by the SmartCheck system. Defects #1–#3 represent the “notched” drill bit, the drill bit with a “blunt edge”, and the drill bit with a “blunt head”, respectively. “Healthy” represents a drill bit without any defects.

Short-time Fourier transform (STFT) was selected as the feature extraction method in this experiment. STFT applies Fourier transform for localization in both the frequency and time domains of time-varying signals [24]. Because of the advantages of frequency-domain analysis of nonstationary signals, STFT is considered a promising approach for analyzing vibration signals under complex conditions or with background noises [15].

A spectrogram takes several FFTs and overlaps them to show how the frequency domain changes with time. It could be a powerful method to describe how a vibration spectrum changes when vibration analysis is performed in a noisy environment.

The STFT process can be represented as follows:

Signal frames are extracted at regular time intervals by a finite window function $w[m]$,

$$x_l[m] = x[m + lH]w[m] \quad (1)$$

where $m \in \{1, 2, \dots, M\}$ is the local time, M is the analysis window length, l is the total number of frames, and H is the hop size. Furthermore, discrete Fourier transform (DFT) was performed on every frame to localize a two-sided spectrum using the following equation:

$$X[k, l] = \frac{1}{M} \sum_{m=1}^M x_l[m] e^{-j2\pi \frac{mk}{K}} \quad (2)$$

where $k \in \{1, 2, \dots, N\}$ is the frequency band index and K is the DFT size. The term $x[k, l]$ represents the STFT of $x[n]$ and corresponds to the local time–frequency behavior of the signal around the time index l and the frequency range k [25]. Figure 5 presents the spectrograms for the different drill bits. In a spectrogram, the color reflects the signal response. In this case, a brighter color reflects a stronger signal response. This feature facilitates the use of 2D CNN algorithms to analyze and learn the various patterns in each health condition. As shown in the figure, most vibration features concentrate in a frequency range lower than 1000 Hz. Therefore, all of the data in the frequency range above 2000 Hz were removed in this experiment. The generated spectrograms were applied directly to the 2D CNN for classification and training.

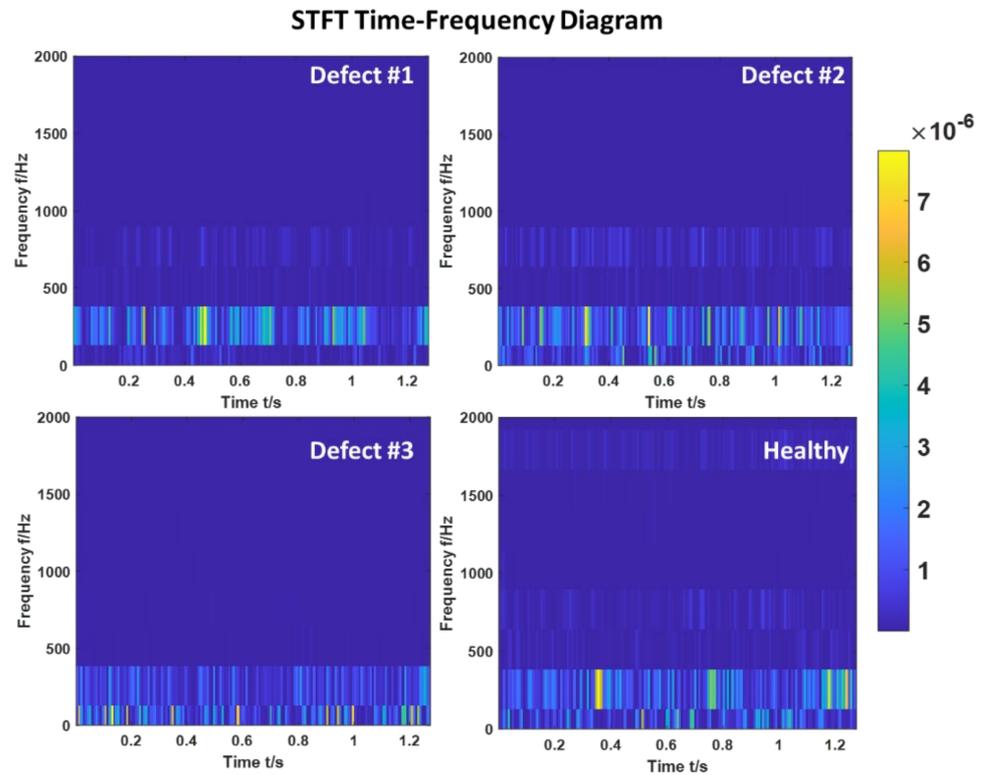


Figure 5. Spectrogram images of the vibration data as the input into the 2D CNN.

Although 2D CNN is a promising method for fault diagnosis applications, it needs a massive amount of calculation to provide a high diagnostic accuracy. Furthermore, 2D CNN requires images as the input to perform the algorithm. It becomes a challenge to apply 2D CNN to real-time vibration analysis. One-dimensional CNN uses 1D data arrays as inputs. It is more suitable for vibration signal analysis. Therefore, 1D CNN was also performed for comparison purposes.

In order to perform the 1D CNN, the collected data were reconstructed to a 1D data structure. Based on Equation (2), the corresponding energy spectral density for the selected frequency band can be presented as follows:

$$X[F1, Tn], X[F2, Tn], X[F2, Tn] \dots X[Fn, Tn] \tag{3}$$

and the spectrogram can be considered as a $F \times T$ matrix. In this experiment, the size of F is 51 and $F \in \{0, 256, 512, 768, 1024, 1280 \dots 12800\}$. In addition, $T = 163$, meaning there are 163 energy spectral density data points for each frequency band. The construction process can be shown as follows:

$$DataArray = [X(F1) + X(F2) + X(F3) \dots X(Fn)] \tag{4}$$

After analyzing the vibration signals, it was found that most of the vibration signals were observed in a frequency range lower than 1300 Hz. Therefore, only a frequency range lower than 1300 Hz was kept during the data reconstruction process. A 1×1536 matrix was created to represent each measurement. The data size was reduced significantly. It led to a reduction in calculation complexity and increased the processing speed. Figure 6 shows the reconstructed data set for different health conditions. The y -axis shows the corresponding signal response to each data point in the data set. The figure presents an obvious difference in the signal response of the different drill bits. For example, in defect #1, there are obvious peaks in the range of the data points from 800 to 1000. On the other hand, the figure shows that most of the peaks are constructed in the first 200 data points for the

healthy drill bit. Defects #2 and #3 show high similarities to each other. The reconstructed data sets were used as the input data in the 1D CNN process for further analysis.

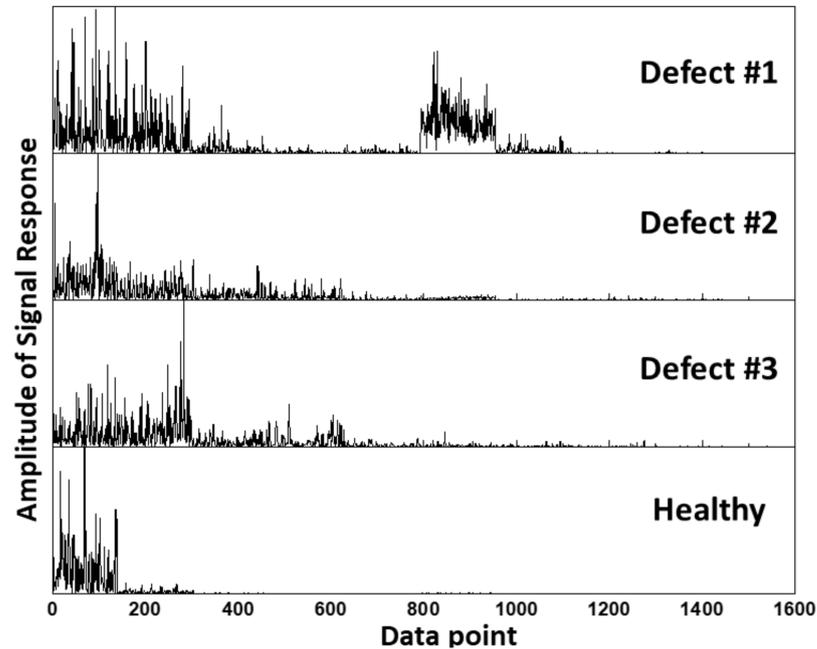


Figure 6. Plot of the reconstructed data sets for the drill bits with different health conditions.

4. Results and Discussion

4.1. CNN Classification Result

In this experiment, two CNN structures were established for defect detection and classification. The structures of both CNN processes are presented in Figure 7. For the 2D CNN, the input data were a set of labeled spectrogram images. The size of the input image was 100×100 pixels. The first convolution layer contained 11 filters with a size of 3×3 . The second convolution layer contained 10 filters with a size of 3×3 . After the convolution layers, a fully connected layer was added. For the 2D CNN, 70% of the labeled data were used to train the network, and 30% of the data were used as the testing samples.

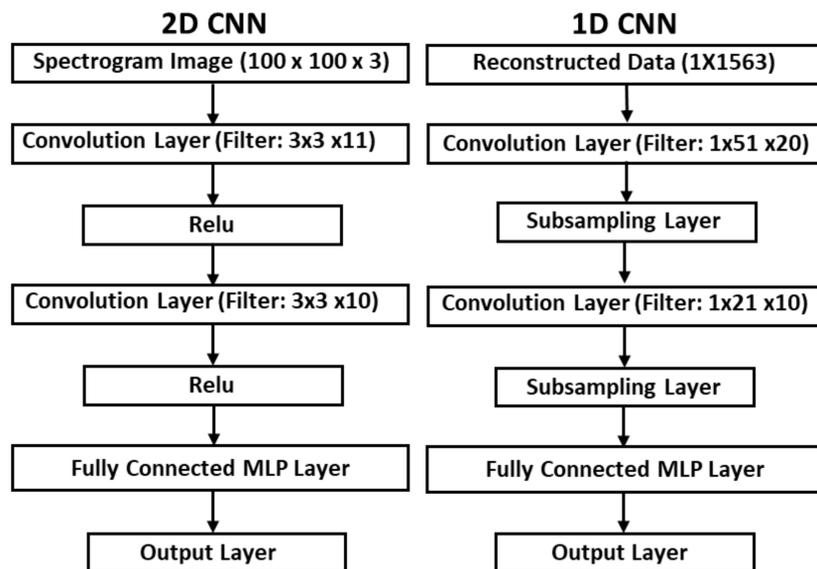


Figure 7. Structures of the 2D CNN and 1D CNN defect diagnosis processes.

For the 1D CNN structure, the first convolution layer contained 20 filters with a kernel size of 51 and was followed by four sub-sampling layers. The second convolution layer contained 10 filters with a kernel size of 21 and two sub-sampling layers. A fully connected layer was added with four classification outputs.

The accuracy of the classification results was calculated based on the equation below:

$$Accuracy = \frac{TP + TN}{TP + TN + FN + FP} \times 100\% \tag{5}$$

where *TP*, *TN*, *FP*, and *FN* represent true positive, true negative, false positive, and false negative, respectively. *TP* indicates that the positive training data are evaluated as positive; *TN* indicates that the negative training data are evaluated as negative; *FP* indicates that the negative training data are evaluated as positive; and *FN* indicates that the positive training data are evaluated as negative.

Figure 8 presents the training and validation results of the drill bit defect diagnosis using the CNNs. Figure 8a shows the confusion matrix of the final learning results. As the confusion matrix shows, the trained model accurately distinguishes defect #1 and the healthy drill bit from the other drill bits. The overall test accuracy is 97.6%. Such high accuracy is considered to be due to the unique vibration patterns caused by the notched edge in the drill bit.

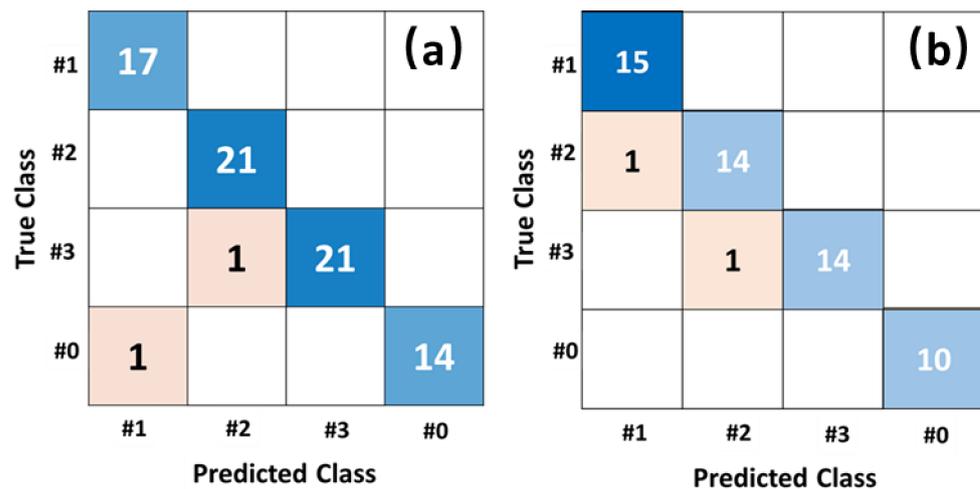


Figure 8. (a) The confusion matrix of the 2D CNN and (b) the confusion matrix of the 1D CNN.

Drilling is one of the typical rotating processes; an inherent vibration can be observed during the operation. When drilling with a notched drill bit, the notch will create special vibration patterns along the turning blade. Such vibration patterns should be able to be distinguished relatively easily [26]. Therefore, it can be observed that defect #1 has a high diagnostic accuracy. For the blunt edge (defect #2) and the blunt head (defect #3), both types of defects increase the drilling resistance during the operation. These defects can be considered as adding loadings to the motor. Several studies have reported how vibration patterns are affected by adding loadings to a motor [27,28]. However, compared to adding loadings to a motor directly, the loadings added to a drill bit with a blunt blade should be relatively small. Therefore, it causes a diagnostic error as reflected in the confusion matrix.

In the 1D CNN, the overall accuracy is about 96.4%. The confusion matrix is shown in Figure 8b. The 1D CNN also shows a high classification accuracy. The classification results of the 1D CNN are similar to those of the 2D CNN. Similar to the 2D CNN results, there is a misjudgment in defect #2 and defect #3. This indicates that the vibration patterns are similar between the blunt edge and blunt head defects. The classification accuracy of each class is 100% for defect #1 and the healthy drill bit, and 93% for defects #2 and #3.

4.2. t-SNE Analysis for 1D CNN

For further analysis, a t-SNE analysis was performed. t-SNE is a kind of visualization algorithm which transforms the high-dimensional Euclidean distance between data points into a conditional probability representing similarity [29]. This analysis maps high-dimensional data into low-dimensional data while keeping the features of the distance between data points. It visualizes if a data set has good divisibility by mapping high-dimensional data to a 2D data set. t-SNE technology is also employed to visualize the learned features of a classification accuracy analysis. The obtained results are presented in Figure 9.

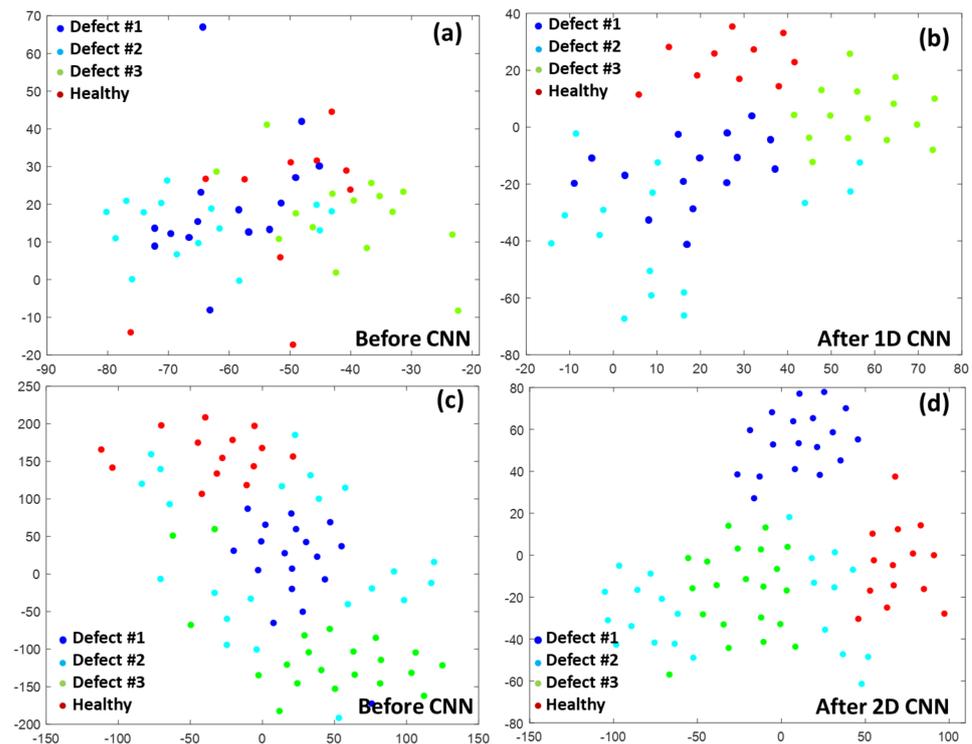


Figure 9. The t-SNE visualization of features for (a) 1D CNN non-processed data, (b) 1D CNN processed test data, (c) 2D CNN non-processed data, and (d) 2D CNN processed test data.

Figure 9 shows the t-SNE results for the test data set before and after applying the 1D convolution and 2D convolution process receptivity. Figure 9a,c present the original data set. Figure 9b,d show the features obtained after the convolutional layers. Compared to the original data set, it clearly shows that all the data points have been gradually separated. In Figure 9b,d, the “healthy” samples are separated from the others clearly. Some overlapping or mixed data points can be observed between defect #1 and defect #2. The results match the classification results shown on the confusion matrix, where the “healthy” sample has 100% accuracy, and there are some classification errors in defect #2 and defect #3.

4.3. Computational Time

In this experiment, MATLAB, R2021b version was used. The PC used for all calculations was equipped with the 11th Gen Intel(R) Core (TM) i5-1135G7 @ 2.4 GHz, with 8GB RAM. For the 2D CNN, 206 s was used for 200 epochs, and approximately 300 s was used for 300 epochs. On the other hand, for the 1D CNN, the overall computational time was around 147 s for 300 epochs. As a result, the processing time for the 1D CNN process was shorter than the 2D CNN process due to the smaller data size.

Table 1 shows the diagnostic accuracy under different epoch numbers for the 2D CNN and the 1D CNN classification processes. For the 1D CNN, classification accuracy increased

rapidly for the initial 100 epochs. After the accuracy reached 80%, the increase in accuracy slowed and gradually reached 96% at 300 epochs. For the 2D CNN, the accuracy stabilized at 97.33% after 200 epochs. As mentioned above, it took approximately 147s for the 1D CNN to run 300 epochs and 206 s for the 2D CNN to run 200 epochs. These results indicate that the 2D CNN has a better classification ability. Although it takes more iterations for the 1D CNN to reach a high classification accuracy, the computational complexity for each iteration is lower than the 2D CNN. In the 2D CNN, the input image size is 100×100 pixels with three layers. It means that there are $10,000 \times 3$ data points for a single measurement. However, for the 1D CNN, the input data array only contains 1536 data points for each measurement. As a result, the overall time cost for the 1D CNN is less than that for the 2D CNN. Additionally, the hardware requirements for computation are much less than the 2D CNN. This property will be beneficial when performing real-time data analysis, and it also provides the possibility for system integration.

Table 1. Diagnostic accuracy with different epoch numbers for the 1D CNN and the 2D CNN.

Epoch Number	1D CNN		2D CNN	
	Accuracy (%)	Time (s)	Accuracy (%)	Time (s)
1	27.7	0.42	29.33	1
50	60	22.1	28.0	52
100	81.8	30.3	57.33	96
150	81.8	57.7	94.67	156
200	87.2	93.8	97.33	206
250	90.9	121.3	97.33	252
300	96.6	147.5	97.33	310

5. Conclusions

This study applied two CNN processes to drill bit defect diagnosis. In the results, both the 2D CNN and the 1D CNN provide high diagnostic accuracy of over 96%. Short-time Fourier transform was performed to convert the time-domain vibration signals into a 2D image to perform the 2D CNN process. The generated spectrum images were directly used as inputs to perform the 2D CNN. In order to perform the 1D CNN, the vibration signals were reconstructed based on the STFT spectrogram to obtain a 1D data set. The reconstructed data set contains 1536 data points. Compared to the raw data set, which contains 32,770 data points, the reconstruction process significantly reduces the data size while keeping the necessary features.

By comparing the two CNN algorithms, we found that the 2D CNN process achieves a higher diagnosis accuracy with a longer computational time. For the 1D CNN, the classification accuracy can be achieved at a similar level to that of the 2D CNN; however, the computational time of the 1D CNN is shorter than the 2D CNN due to its low computational complexity. This process reduces the hardware requirements for computation and can be beneficial to system integration and real-time analysis. It is worth mentioning that both CNN processes provide high accuracy in distinguishing “healthy” conditions from damaged conditions. In other words, the processes can detect whether the drill bits are damaged. For fault diagnosis purposes, the overall diagnostic accuracy of the 1D CNN is lower than that of the 2D CNN, but it does provide a high calculation speed and an accurate diagnosis.

Author Contributions: Conceptualization, Y.Y. and W.Z.; methodology, Y.Y., Q.L. and W.Z.; original draft preparation, Y.Y.; review and editing, Q.L., B.S.H. and W.Z.; supervision, W.Z.; project administration, B.S.H. and W.Z. All authors have read and agreed to the published version of the manuscript.

Funding: This research is funded by the Agency for Science, Technology, and Research (A*STAR) and Schaeffler for project entitled “Condition Monitoring and Fault Prognosis of Robot Joints based on Fiber Optic Sensing and Deep Learning” through grants #020956-00003 and #021003-00003.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: This research is supported by the Agency for Science, Technology, and Research (A*STAR) under its IAF-ICP Programme I2001E0067 and the Schaeffler Hub for Advanced Research at NTU.

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

Abbreviations

The following abbreviations are used in this manuscript:

AI	Artificial Intelligence
CNN	Convolutional Neural Network
DBN	Deep Belief Network
DFT	Discrete Fourier Transform
EMD	Empirical Mode Decomposition
MLP	Multi-Layer Perceptron
RGB	Red, Blue, and Green
RMS	Root Mean Square
STFT	Short-Time Fourier transform
SVM	Support Vector Machine
t-SNE	t-distributed Stochastic Neighbor Embedding
VMD	Variational Modal Decomposition
WPE	Wavelet Packet Energy

References

1. Vununu, C.; Moon, K.-S.; Lee, S.-H.; Kwon, K.-R. A deep feature learning method for drill bits monitoring using the spectral analysis of the acoustic signals. *Sensors* **2018**, *18*, 2634. [[CrossRef](#)] [[PubMed](#)]
2. Dai, Y.; Zhu, K. A machine vision system for micro-milling tool condition monitoring. *Precis. Eng.* **2018**, *52*, 183–191. [[CrossRef](#)]
3. Bhuiyan, M.; Choudhury, I.A.; Dahari, M.; Nukman, Y.; Dawal, S. Application of acoustic emission sensor to investigate the frequency of tool wear and plastic deformation in tool condition monitoring. *Measurement* **2016**, *92*, 208–217. [[CrossRef](#)]
4. Asr, M.Y.; Etefagh, M.M.; Hassannejad, R.; Razavi, S.N. Diagnosis of combined faults in Rotary Machinery by Non-Naive Bayesian approach. *Mech. Syst. Signal Process.* **2017**, *85*, 56–70. [[CrossRef](#)]
5. Georgoulas, G.; Karvelis, P.; Loutas, T.; Stylios, C.D. Rolling element bearings diagnostics using the Symbolic Aggregate approxImation. *Mech. Syst. Signal Process.* **2015**, *60*, 229–242. [[CrossRef](#)]
6. Xiong, Q.; Zhang, W.; Lu, T.; Mei, G.; Liang, S. A fault diagnosis method for rolling bearings based on feature fusion of multifractal detrended fluctuation analysis and alpha stable distribution. *Shock Vib.* **2016**, *2016*, 1232893. [[CrossRef](#)]
7. Zhang, M.; Jiang, Z.; Feng, K. Research on variational mode decomposition in rolling bearings fault diagnosis of the multistage centrifugal pump. *Mech. Syst. Signal Process.* **2017**, *93*, 460–493. [[CrossRef](#)]
8. Zhang, J.; Yi, S.; Liang, G.; Hongli, G.; Xin, H.; Hongliang, S. A new bearing fault diagnosis method based on modified convolutional neural networks. *Chin. J. Aeronaut.* **2020**, *33*, 439–447. [[CrossRef](#)]
9. Chuya-Sumba, J.; Alonso-Valerdi, L.M.; Ibarra-Zarate, D.I. Deep-Learning Method Based on 1D Convolutional Neural Network for Intelligent Fault Diagnosis of Rotating Machines. *Appl. Sci.* **2022**, *12*, 2158. [[CrossRef](#)]
10. Neupane, D.; Kim, Y.; Seok, J.; Hong, J. CNN-Based Fault Detection for Smart Manufacturing. *Appl. Sci.* **2021**, *11*, 11732. [[CrossRef](#)]
11. Wang, H.; Li, S.; Song, L.; Cui, L.; Wang, P. An enhanced intelligent diagnosis method based on multi-sensor image fusion via improved deep learning network. *IEEE Trans. Instrum. Meas.* **2019**, *69*, 2648–2657. [[CrossRef](#)]
12. Hoang, D.-T.; Kang, H.-J. Rolling element bearing fault diagnosis using convolutional neural network and vibration image. *Cogn. Syst. Res.* **2019**, *53*, 42–50. [[CrossRef](#)]
13. Janssens, O.; Slavkovikj, V.; Vervisch, B.; Stockman, K.; Loccufer, M.; Verstockt, S.; Van de Walle, R.; Van Hoecke, S. Convolutional neural network based fault detection for rotating machinery. *J. Sound Vib.* **2016**, *377*, 331–345. [[CrossRef](#)]
14. Ding, X.; He, Q. Energy-fluctuated multiscale feature learning with deep convnet for intelligent spindle bearing fault diagnosis. *IEEE Trans. Instrum. Meas.* **2017**, *66*, 1926–1935. [[CrossRef](#)]
15. Verstraete, D.; Ferrada, A.; Droguett, E.L.; Meruane, V.; Modarres, M. Deep learning enabled fault diagnosis using time-frequency image analysis of rolling element bearings. *Shock Vib.* **2017**, *2017*, 5067651. [[CrossRef](#)]

16. Eren, L. Bearing fault detection by one-dimensional convolutional neural networks. *Math. Probl. Eng.* **2017**, *2017*, 8617315. [[CrossRef](#)]
17. Eren, L.; Ince, T.; Kiranyaz, S. A generic intelligent bearing fault diagnosis system using compact adaptive 1D CNN classifier. *J. Signal Process. Syst.* **2019**, *91*, 179–189. [[CrossRef](#)]
18. Ince, T.; Kiranyaz, S.; Eren, L.; Askar, M.; Gabbouj, M. Real-time motor fault detection by 1-D convolutional neural networks. *IEEE Trans. Ind. Electron.* **2016**, *63*, 7067–7075. [[CrossRef](#)]
19. Kao, I.-H.; Wang, W.-J.; Lai, Y.-H.; Perng, J.-W. Analysis of permanent magnet synchronous motor fault diagnosis based on learning. *IEEE Trans. Instrum. Meas.* **2018**, *68*, 310–324. [[CrossRef](#)]
20. Yari, M.; Bagherpour, R.; Khoshouei, M. Developing a novel model for predicting geomechanical features of carbonate rocks based on acoustic frequency processing during drilling. *Bull. Eng. Geol. Environ.* **2019**, *78*, 1747–1759. [[CrossRef](#)]
21. Yari, M.; Bagherpour, R. Investigating an innovative model for dimensional sedimentary rocks characterization using acoustic frequencies analysis during drilling. *Rud. Geološko Naft. Zb.* **2018**, *33*, 17–25. [[CrossRef](#)]
22. Khoshouei, M.; Bagherpour, R.; Jalalian, M.H.; Yari, M. Investigating the acoustic signs of different rock types based on the values of acoustic signal RMS. *Rud. Geološko Naft. Zb.* **2020**, *35*, 29–38. [[CrossRef](#)]
23. Jeong, M.-J.; Lee, S.-W.; Jang, W.-K.; Kim, H.-J.; Seo, Y.-H.; Kim, B.-H. Prediction of drill bit breakage using an infrared sensor. *Sensors* **2021**, *21*, 2808. [[CrossRef](#)] [[PubMed](#)]
24. Benesty, J.; Sondhi, M.M.; Huang, Y. *Springer Handbook of Speech Processing*; Springer: Berlin/Heidelberg, Germany, 2008; Volume 1.
25. Zhivomirov, H. On the Development of STFT-analysis and ISTFT-synthesis Routines and their Practical Implementation. *TEM J.* **2019**, *8*, 56–64.
26. Li, B.; Chow, M.-Y.; Tipsuwan, Y.; Hung, J.C. Neural-network-based motor rolling bearing fault diagnosis. *IEEE Trans. Ind. Electron.* **2000**, *47*, 1060–1069. [[CrossRef](#)]
27. Rodríguez, P.V.J.; Negrea, M.; Arkkio, A. A simplified scheme for induction motor condition monitoring. *Mech. Syst. Signal Process.* **2008**, *22*, 1216–1236. [[CrossRef](#)]
28. Goundar, S.; Pillai, M.; Mamun, K.; Islam, F.; Deo, R. Real time condition monitoring system for industrial motors. In Proceedings of the 2015 2nd Asia-Pacific World Congress on Computer Science and Engineering (APWC on CSE), Nadi, Fiji, 2–4 December 2015; pp. 1–9.
29. Wang, X.; Mao, D.; Li, X. Bearing fault diagnosis based on vibro-acoustic data fusion and 1D-CNN network. *Measurement* **2021**, *173*, 108518. [[CrossRef](#)]