Learning Convolution Neural Network with Shift Pitching based Data Augmentation for Vibration Analysis

M. F. M. Esa[1] , N. H. Mustaffa[1], H. Omar[2], N. H. M Radzi[1], and Sallehuddin,R[1],

[1]School of Computing, Faculty of Engineering, Universiti Teknologi Malaysia 81310 Johor Bahru, Johor, Malaysia
[2]Centre for Foundation Studies, City University Malaysia, 46100 Petaling Jaya, Selangor, Malaysia

Email: pkharmony@yahoo.com

Abstract. Data augmentation is a common approach that been implemented in order to increase the training data quantity for Convolutional Neural Networks in signal processing, image recognition and speech recognition. However, the conventional data augmentation methods usually implement the window slicing and overlap window slicing methods in the bearing fault analysis. Meanwhile, the audio deformation approach such as time stretching and pitch shifting methods have been commonly used as data augmentation approach in speech recognition. Thus, this paper proposed a data augmentation based on shift pitching technique for the vibration signal. The relationship between the audio and the vibration signal is evaluated for a bearing fault analysis using Convolution Neural Networks. The new dataset produce by the data augmentation is used to increase the number of training dataset and to improve the Convolutional Neural Networks training performance. The result shows that the shift pitching based data augmentation method able to achieve higher training accuracy compared to the window sliding data augmentation. The combinations of all ratio pitch obtain 93% accuracy whilst the accuracy for a single rate pitch are between 81% to 91%.Thus, the proposed method is competent and able to improve the performance of bearing fault classification

1. Introduction

In a data acquisition system for vibration analysis, an accelerometer sensor is used to capture a vibration signal from the vibration and noisy environment. There is various type of accelerometer with different specification available in market which affected the quality of vibration signal. In addition, accelerators with the same series output can also produce different signal quality [1]. Another factor is included point position, device orientation, distance vibration source and environment noise [2-3].

After signal preprocessing stage, all collected signal will be analysis using different technique. The combination of indicators from the different signal will determine the level of bearing severity and subsequent action necessary. However, the whole process to interpret various signal data and to make the conclusion in vibration analysis is requires the expert in field [4]. To overcome the mentioned situation, the application of signal processing and machine deep learning is promise the solution.

Recently, Convolutional Neural Networks (CNN) from a type of deep learning [5-6] has received a lot of interest in vibration analysis due to its major advantages in extracting complex feature from signal automatically [7-9]. This capability is significantly contributing the reduction of expert effort in vibration analysis process. Moreover, CNN is the noise tolerance [3]. There are many CNN application proven this advantages in problem of the signal with noise.

Meanwhile, the disadvantage of CNN is its structure model consisted of hyper-parameters and requires a large dataset for training to be sufficient in learning the problem. This problem is compounded with the problem on the limitation collection of vibration signal data for the defect bearing due to of depending to the device specification, laborious and expensive cost. This problem seriously devastating effect to the performance and threatening CNN to achieve good generalization model and prevent over fitting problem. Over fitting problem is happening whereas a CNN model can train the data very well but the testing with other data that exclude in the training data show the low result significantly.

To overcome the over fitting problem, data augmentation (DA) method has extensively been proposed in literature [10]. DA is a method to increase the quantity of data by create potential latent variables or hidden data in population data [11]. Therefore, with these abilities, DA has worked excellently for training CNN model in signal processing, image recognition [12] and speech recognition [13-14].

Several DA methods from image processing has been implemented into signal processing such as window slicing (WS) [14-15] and overlapping windows methods [16]. Meanwhile, an audio deformation method is commonly used as a data augmentation technique in the speech recognition, Audio deformation method artificially converted the raw audio signal data to a different tone of sound but has ability to preserve the semantic meaning of the signal [17]. Pitch shifting, dynamic range compression [10], time stretching and speed perturbation [13] are the example of the audio deformation methods that has been used for data augmentation. The audio deformation using speed and amplify transformation approaches have been successfully implemented in CNN for classification of vibration signal [17]

Pitch is a sound feature that allows their interpreting a frequency-related scale in audio data such as in human voice [13]. Inspire from nature human voice that have inconsistently pitch cause many factor like gender and age, we analogy this phenomena with vibration signal that captured using different sensor and point position. For this theory, the meaning of indicator is same even the signal pitch is different.

Thus, this paper proposed the shift pitching for vibration signal to evaluate relationship between audio and vibration signal using bearing fault analysis using CNN. Furthermore, DA based on shift pitch are proposed as approach to increase the number of training dataset for vibration signal and to evaluate the performance of CNN's based on various ratio of pitch.

2. Methodology

The vibration analysis in this study is conducted by several steps starting with the preparation of the raw vibration signal dataset. The dataset is separated into 3 group to be used for training, validation and testing in CNN. The proposed DA only been applied for the training data. Next, the long length of 1-D time series vibration signal is segmented into a shorter length than the original times series data. Each segmented data will transform to frequency domain format using FFT function. Furthermore, CNN model will be train, validated and testing to produce result. The flowchart of the proposed shift pitching method is shown in Figure 1.
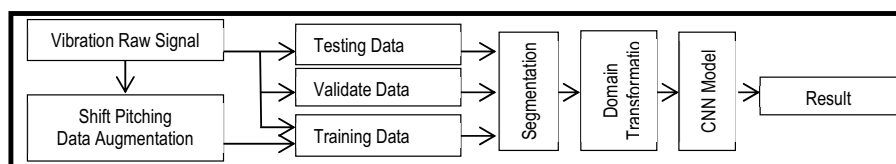
**Figure 1.** Flow chart of the proposed shift pitching method

2.1. CNN Model

CNN is a multi-layer artificial neural network (ANN) that is structured by kernel layers and ending with classification layer [14]. The kernel layer is used to abstract input data representation, that is usually consist of three basic operations. The first layer is a convolutional layer followed by the activation

function (RELU layer) and the pooling layer. The classification layer at the end of the structure is composed of fully connected and output layers.

2.2. Shift Pitching based Data Augmentation
A data augmented based on shift pitching technique is proposed to increase the CNN training dataset for vibration analysis. Hence, the issue of limitation of signal data can be resolved with the capability of this approach where it still able to preserve the semantic meaning eve after the signal has been modified. Matlab® Software is used as a tool to transform the raw signal data. The shift pitching method produce a variant of signal at frequency angle (x axis) of signal spectrum graph. Figure 2 shows the comparison of different signal pattern at low, original and high rate of pitch.
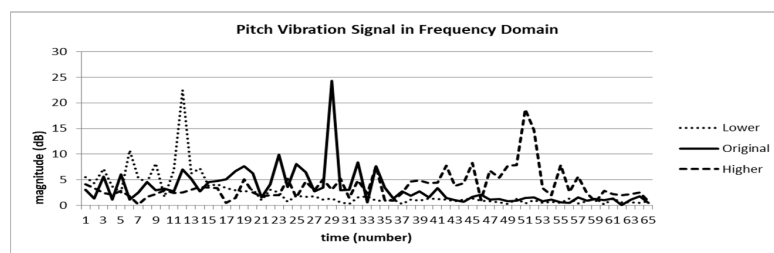
Figure 2. Same raw signal pattern with different rate pitch in frequency domain.

2.3. Experimental Evaluation
The experimental vibration signal dataset collected by Case Western Reserve University (CWRU) [18] are used to validate the performance of proposed method. The dataset is consisting of 10 condition for four motor load, two frequency, three speed rate and four size of bearing. This properties of the dataset used in the analysis are as follows: 1 hp load, 48khz frequency, 1772 rpm and 10 vibration signal that represents a good condition bearing and 9 type of defects. The signal is executed from a raw signal and the 8 additional signals are produces using a pitch shifting method as an augmentation technique. The pitch value are scaled to a ratio of 20, 10, 110,120,130,140 and 150 in percent for the analysis for each transformation and the analysis also evaluated the performance of the combination for all pitch ratio

Then the fix window technique is implemented using the 2048 size window to produce the 200 sub signal for each signal. The generated signals data are divided into 70% for training, 15% for validation and 15% for testing. The analysis is run for 20 times for each experiment using Python and Matlab2018b The CNN structure and the training parameters setting used for the experiments are shown in Table 1.

Table 1. Structures of the CNN, training and variables of parameters.

Layer	Type	Variables and Dimensions	Training Parameters
1	Convolution	FN = 10; FS = 256; B = 10	• SGD minibatch size = 20
2	Pooling	S = 2	• Weight decay = 0.04
3	Convolution	FN = 15; FS = 32; B = 15	• Momentum = 0.5
4	Pooling	S = 2	• Epochs = 25
5	Convolution	FN = 15; FS = 16; B = 15	• Reduction of learning rate after each ten epoch.
6	Hidden layer	Relu activation function	
7	Softmax	10 outputs	• Testing sample rate = 15%

SDG = stochastic gradient decent; No overlapping of convolutional window and no padding. FN = filter number, FS = filter size; B = bias; S = down-sampling rate;

3. Results and Discussion

Figure. 3 shows the result on CNN training accuracy for a window sliding method(label as 0), shift pitching based DA method for 7 pitch ratio scale (20,20,10,120,130,140 and 150) and with combination of all pitch ratio values.
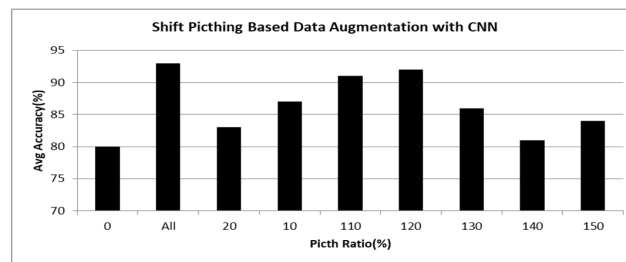
Figure 3. Comparison of CNN training accuracy with different ratio pitch

The result shows that the pitch shifting method is competent to analysis bearing fault compared to the window sliding method. The combinations of all ratio pitch obtain the highest accuracy with 93% accuracy and the accuracy for a single rate pitch are varied between 81% to 91%. The result also shows that the scale of pitch ratio also effect the performance of CNN training accuracy where the accuracy is increase as the pitch ratio increases. However, too large pitch ratio scale compared with the original data will lower the accuracy.

This result is significantly correlated with original training dataset that less variant in pitch, meanwhile the proposed method able to create a pattern more close like in testing data. Hence, the proposed approach is superior to the window sliding data augmentation method due to its ability to generate the additional and variant dataset. This contribution is from the efficiency of introducing additional dataset for vibration analysis using proposed approach.

4. Conclusion

The shift pitching based data augmentation method for training CNN is proposed for vibration analysis in bearing fault diagnosis. The idea inspires from different pitch of different human voice recite the same sentence. Shift pitching is implemented in this paper by combining the raw signals data in order increase the training data for CNN. The experiment is conducted using bearing fault diagnosis to evaluate the performance of the proposed shift pitching method. The result shows that the proposed method ableto obtain higher accuracy compared with the window sliding method. The proposed CNN architecture, in combination with shifting pitch DA, produces promising performance for bearing fault classification. Furthermore, the proposed approach is capable to apply for different type of signals in signal processing that have limitation of data sources. The fusion of various bearing and sensor specification will be considered in the future research.

References

[1]   Slaven J E, Andrew M E, Violanti J M, Burchfiel C M and Vila B J 2006. *Physiological Measurement* **27** 413.

[2]   Gordon S D, Tiller B, Windmill J F C, Krungner R and Narins P M 2019. *J Comp Physiol A* **205** 783–91

[3]   Zhang W, Peng G, Li C, Chen Y and Zhang Z 2017 *Sensors* **17** 425.

[4]   Ebersbach S. and Peng Z 2008 *Expert Syst. Appl.*, **34** 291-9

[5]   Lemley J, Shabab B and Peter C 2017 *IEEE Access* **5** 5858-69.

[6]   Russakovsky O, Deng J, Su H, Krause J, Satheesh S, Ma S,Huang Z, Karpathy A, Khosla A, Bernstein M, Alexander C B and Fei-Fei L 2015 *Int J Comput Vis* **115** 211–52

[7]   Tandon N and Choudhury A 1999 *Tribology International* **32** 469-80

[8]   Cui Z, Wenlin C and Yixin C 2016 Multi-scale CNNs for time series classification *Preprint* arXiv:1603.06995

[9]   Lu W, Liang B, Cheng Y, Meng D, Yang J and Zhang T 2016 *IEEE Trans. Ind. Electron*. **64** 2296-2305

[10]   Salamon, Justin and Bello J P 2017 *IEEE Signal Processing Letters* **24** 279-83.

[11]   Van Dyk D A and Meng X L 2001 *J Comput. Graph Stat.* **10** 1-50

[12]   Howard A 2013 Some improvements on deep CNN based image classification *Preprint* arXiv:1312.5402,

[13]   Ko T, Peddinti V, Povey D and Khudanpur S 2015 1*6th Annual Conf. of the Int Speech Communication Assocciation (Dresden)*

[14]   Rebai I, BenAyed Y, Mahdi W and Lorré J P 2017 *Procedia Computer Science* **112** 316-22.

[15]   Le Guennec A, Malinowski S and Tavenard R 2016 *ECML/PKDD Workshop on Advanced Analytics and Learning on Temporal Data* (*Riva Del Garda, Italy*)

[16]   Park C S, Choi Y C and Kim Y H 2013 *Mech. Syst. Signal Process* **38** 534-48

[17]   Esa M F M, Mustaffa N H, Radzi N H M and Sallehuddin R 2019 *IOP Conf. Series: Materials Science and Engineering* **551** 012066.

[18]   Loparo K 2012 *Case Western Reserve University Bearing Data Centre* Online: http://csegroups.case.edu/ bearingdatacenter/pages/download-data-file