A Multi-task Deep Learning Model for Rolling-Element Bearing Diagnostics

Ahmed Zakariae Hinchi[1], Mohamed Tkiouat[2]

[1,2]*Laboratory for Applied Mathematics (LERMA), Mohammadia School of engineering, Mohammed V University, Rabat, Morocco*

ahm.zak.hin@gmail.com
tkiouat@emi.ac.ma

ABSTRACT

The rolling element bearing is the predominant source of failures in rotating machinery. Therefore, detecting the corresponding faults, predicting their locations and measuring their severity is of immense importance. Classical intelligent diagnostics approaches rely on feature extraction methods followed by a classification model. Recently, deep learning models have improved the fault classification accuracy by learning a suitable representation directly from the raw sensor data. In this work, we present a novel multi-task deep convolutional neural network trained end-to-end on raw vibration data to learn a shared representation for fault isolation and fault size evaluation. The proposed model architecture is constructed by stacking blocks of convolution layers, pooling layers, and batch normalization layers followed by a regression head and a classification head. Extensive experiments show that the proposed approach produces a superior performance to other existing methods and generalizes well to fault sizes not present in the training set.

1. INTRODUCTION

As the complexity of modern machines rises, the degradation of each component increasingly influences the failure probability of other interconnected parts. This fact, along with the growing challenges of cost and quality in the contemporary industry prompted the change in maintenance policies from corrective to condition-based strategies.

Condition-based maintenance recommends maintenance tasks based on the real-time collected sensor data (Jardine et al., 2006). Moreover, this maintenance program allows faster operations through automated fault isolation and diagnostics, as well as the avoidance of unnecessary maintenance tasks and the anticipation of mandatory ones through prognostics.

In rotating machinery, the rolling element bearing is the most vital element and accounts for the majority of failures (R.L. Winder and W.E. Littmann 1997). Moreover, various techniques are used for the monitoring of this component; from oil debris or motor current signature analysis to the examination of acoustic emissions or more frequently vibrations. Consequently, there are a plethora of studies concerning the intelligent diagnostics of rolling element earing using vibration data.

Traditionally, data-based diagnostics of rolling element bearing is divided into two complementary steps: mapping the high dimensional vibration data to the feature space by extracting, selecting, and combining various features; and classifying the chosen features to the right machine condition using pattern recognition techniques such as support vector machines, Bayes Classifier, k-nearest neighbor, Decision Trees, Random Forests, and neural networks. In this framework, the feature extraction step is the most important one and thus received a notable interest from the research community (Rai and Upadhyay, 2016); early works focused on extracting statistical parameters from the time domains such as kurtosis, root mean square value, or the crest factor (Samanta and Al-balushi, 2003). Other studies derived parameters from the frequency representation or the time-frequency representation using techniques such as the Fast Fourier transform or the Short-time Fourier transform. Moreover, methods such as wavelet transform, empirical mode decomposition, chaos theory, and spectral kurtosis are extensively studied and widely applied (Chen et al., 2016; Lei et al., 2013; Wang et al., 2016; Yan et al., 2014; Yang et al., 2007).

Since 2006, various deep learning models reached state of the art performance in different artificial intelligence applications (LeCun et al., 2015). This performance is attributed to their end-to-end approach and feature-learning paradigm. A deep learning model contains a hierarchy of stacked learning layers each depicting a representation level; i.e., the first layer represents the raw data, and the final layer describes the chosen output. Subsequently, different studies using deep learning for bearing diagnostics have been proposed in recent years. The first works used stacked restricted Boltzmann machines to construct deep belief networks for the hierarchical representation ((Shao et al., 2015), (Deng et al., 2016) and (Tao et al., 2016) (Gan et al., 2016) (He et al., 2017)). Meanwhile, other studies used variant of autoencoders for the end-to-end diagnostics: (Jia et al., 2016) trained a deep neural network by stacking autoencoders on the frequency spectra. Following that, (Guo et al., 2017) proposed a stacked denoising autoencoder to generalize to noisy operating conditions. Moreover, (Shao et al., 2017) presented a deep autoencoder trained greedily on a maximum cross entropy loss function and optimized its architecture using the artificial fish swarm algorithm. Finally, (Jia et al., 2018) proposed a novel normalized sparse autoencoder.

Motivated by the success of deep convolutional neural networks in image and speech recognition, few studies used these efficient architectures for bearing diagnostics; (Guo et al., 2016) proposed a hierarchical model where first the fault is isolated, then the fault size is evaluated in the second step. The authors chose the leNET architecture (a standard architecture for image recognition) and transformed the one-dimensional vibration signal accordingly into a two-dimensional grid. Moreover, (Shao et al., 2018a, 2018b) proposed a convolutional deep belief network where the vibration data is first compressed to increase efficiency. (Li et al., 2017) proposed an ensemble deep convolutional neural network where the decisions are taken by evidence fusion using an improved Dempster–Shafer theory.(Zhang et al., 2017) proposed a deep convolutional network with large first-layer kernels and small kernels in the following layers. This model can achieve high accuracy in a noisy environment and can be adapted to different loading conditions by calculating their statistics. In (Zhang et al., 2018), the authors improved the model by adopting a new training methodology (small batch sizes and a varying dropout rate).The results showed that the new model generalizes well to noisy environments and different loading and operating conditions without knowing their statistics a priori.

Nevertheless, there are apparent limitations with the aforementioned models. First, most of the reviewed approaches represent bearing diagnostics as a classification problem where each combination of fault type and fault size form a category. While this representation is validated to satisfying results in experimental data where fault sizes are seeded to discrete ones known a priori, it fails in practice as the fault size rises continuously with the degradation of the bearing. This fact violates the closed-world assumption and possibly confuses the separating hyperplanes. Second, while some studies adopt a hierarchical approach separating fault isolation and severity assessment, most treat the fault size assessment as a classification problem. Furthermore, the two-step framework requires first training a model to detect and isolate the different faults, then train a model for each fault to assess its size. Thus, losing shared information about fault severity assessment.

The contributions of the proposed approach are summarized as follows. Firstly, the proposed method can achieve perfect diagnostics performance without manual feature extraction. Secondly, the proposed model learn a shared representation for fault isolation and fault size evaluation and can generalize to fault sizes out of the training set.

The remainder of this work is organized as follows. Section 2 presents the model and its training framework. Section 3 investigates the effectiveness of the proposed approach for rolling element bearing fault diagnostics. Finally, Section 4 provides the conclusion.

2. METHODOLOGY

2.1. An introduction to convolutional neural networks

Typically, a deep convolutional model uses a hierarchy of convolutional and pooling layer followed by a fully connected layer(s) to process the input data and estimate the output respectively.
The convolutional layer extracts the feature map from the layer below using a series of trainable kernels called filters. Each filter convolves with its input then pass through an activation function. The convolutional layer determines the results of a convolution operation with input X, a bank of trainable kernels K of length d_k, and a bias b followed by a nonlinear activation function σ:

$$(X*K)_i = \sigma(\sum_{m=0}^{d_k-1} K_m \cdot X_{i+m} + b) \qquad (1)$$

It is common to follow the convolutional layer with a pooling layer. This layer is used to reduce the resolution of the feature map typically by employing a local maximum or average operation (max-pooling or average pooling). The pooling layer compresses the feature space, regularizes the deep neural network, and make the feature space more robust to shifts and distortions.
On the other hand, the Batch Normalization (BN) proposed by (Ioffe and Szegedy, 2015) can lead to faster training times and improved regularization by controlling the input distribution across the layers. The first step is to normalize features on the batch axis using the mean and standard deviation of each activation of the mini-batch.
Finally, the output of the last convolutional layer is flattened and connected to a dense layer. A fully connected or a dense layer is a layer where each neuron is connected to each neuron in the next layer through a weight matrix W_f, a bias b_f, and an activation function σ:

$$f_t = \sigma\left(W_f\left[h_{t-1}, x_t\right] + b_f\right) \qquad (2)$$

The model parameters are usually defined using a gradient descent based minimization of the chosen loss function using the backpropagation algorithm.

2.2. The multi-task training approach

In multi-task learning, the model is trained using multiple loss functions each corresponding to a task. By employing a shared representation, the model can exploit the synergy between tasks and boost the performance of each one (Ando and Zhang, 2005). Prior works suggest that multi-task training allows the learning of a more generalizable embedding. For instance, (Gebru et al., 2017) found that multi-task training improved the domain adaptation in a fine-grained classification framework. Similarly, (Ranjan et al., 2017) presented hyper-face: a model for simultaneous face detection, landmarks localization, pose estimation and gender recognition in which the fused representation allows better generalization for each of the tasks mentioned above. In (Hinchi and Tkiouat, 2018), we used an auxiliary loss to improve the estimation of the remaining useful life of a rolling element bearing.

We propose a multi-task training framework for the join fault isolation and severity assessment (figure 1). Subsequently, we train our model with two loss functions: The classification head is trained with a categorical cross entropy loss measuring the difference between the right distribution p and estimated fault diagnostics distribution q:

$$L_{fault_{isolation}} = \sum_{x} \sum_{class} -y_{class} log(\hat{y}_{class}) \quad (3)$$

On the other hand, the regression head is trained using the mean absolute error loss function:

$$L_{fault\_size} = \frac{\sum |y_{size} - \widehat{y_{size}}|}{N} \quad (4)$$

The model is trained end-to-end using the aforementioned loss functions. Specifically, the backpropagation algorithm is used to compute the gradients. The overall gradient is composed of the gradient from the classification loss function plus the one from the regression one multiplied by a discount weight γ=0.25. The optimization is conducted with stochastic gradient descent.

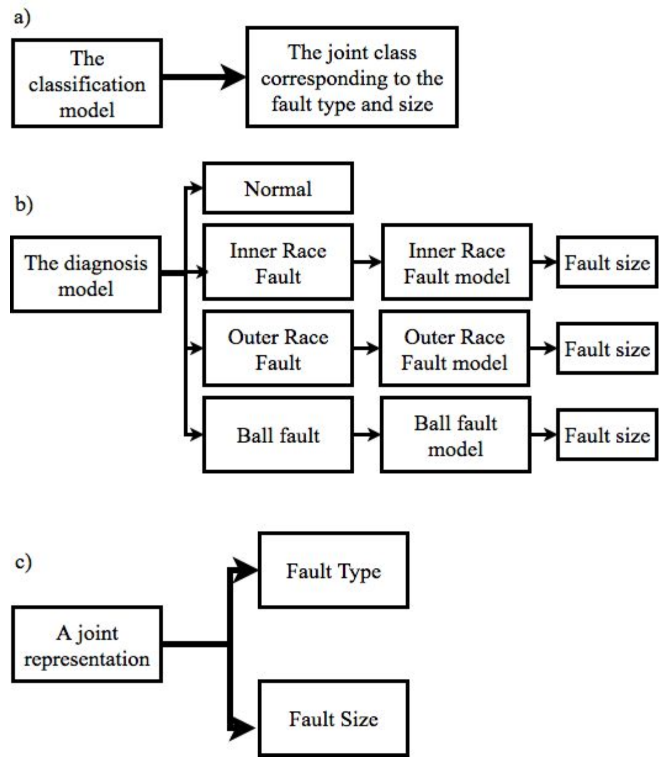

Figure 1. The different frameworks used for fault diagnostics. 1.a: The classification framework. 1.b: The hierarchical framework. 1.c: The proposed joint multi-task framework.

2.3. The model architecture

We propose a convolutional neural network architecture composed of stacked convolutional, batch normalization and pooling layers respectively. Since the fault size in the healthy condition is equal to zero, we multiply the fault size neuron value by one minus the probability of the normal health condition. This simple restriction is visualized in Figure 2. Furthermore, the selection of the hyper-parameters of deep neural networks is an open problem. Various recent works tackled this issue with evolutionary strategies (Young et al., 2015), reinforcement learning (Zoph and Le, 2016), or Bayesian optimization. Although these methods demonstrated good performances, they require a significant computational burden. In this work, we used a simple random search over the space of hyper-parameters (the number of filters and their length and stride, the number of layers…). The stride of the first layer is fixed by limiting the receptive field of each neuron in the dense layer to one period of the input signal. As (Cohen and Shashua, 2016) showed that deep Convolutional rectifier networks are universal with max pooling, we use the rectifier activation $\sigma = \max(0, x)$ after every convolutional layer and the fault size neuron. For the fault detection and isolation layer, we use the softmax loss function. $\sigma(z)_j = \frac{e^{z_j}}{\sum_{class}^{4} e^{z_{class}}}$ for j=1,…4

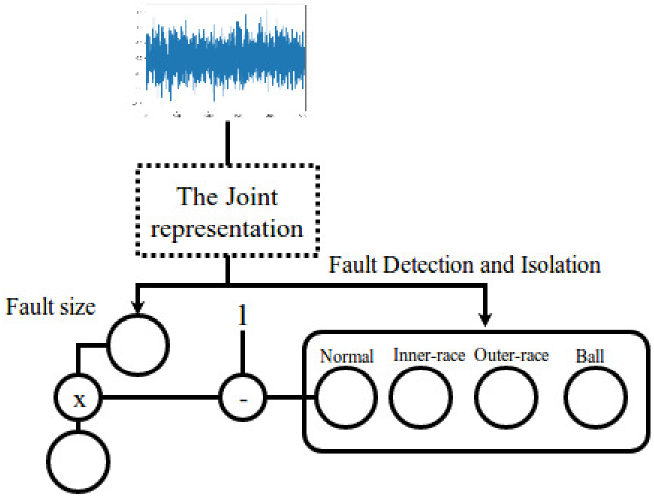

Figure 2. The simplified architecture of the proposed model.

3. VALIDATION OF THE PROPOSED MODEL

In real-world applications, the fault size varies continuously due to the bearing degradation. Thus, it is unrealistic to have every possible fault size present in the experimental set. In the remainder of this section, we will investigate the performance of the three different frameworks from figure 1, and their ability to generalize to new fault sizes.

3.1. Data description

The experiments data are acquired from the Case Western Reserve University (CWRU) Bearing Data center (Loparo, 2005). This dataset contains vibration data sampled at a frequency of 12 kHz under four different operating conditions (0, 1, 2, and 3 horsepower). Along with the normal vibration data, seeded faults are situated in the three different component of the rolling element bearing (ball, inner-race, and outer-race). The seeded faults can have a diameter of 0.007 in., 0.014 in., or 0.021 in.

From every operating condition, we use 7.5 seconds of normal vibration, and 2.5 seconds of the three fault types and the three-fault size respectively with the exception of the 0.014-inch ball fault for the training and validations sets. In addition, the vibration data is divided into samples of 1024 point where the training samples are overlapped with an overlap of 64 for data augmentation. Table 1 describes the detailed dataset for each operating condition.

3.2. Experimental setup

This section investigates the performance of the proposed neural network in the three diagnostics framework explored in figure 1. To limit the effect of the hyper-parameter selection step we use the same architecture for the three model up until the last layer. This architecture is composed of six stacked convolutional layers each followed with a

Table 1. The experimental dataset.

Condition	Training set	Validation set	Test set
Normal	7.5 s / 1359 sample	7.5 s / 87 sample	7.5 s / 87 sample
Inner-race 0.007 in	2.5 s / 453 sample	2.5 s /29 sample	2.5 s /29 sample
Inner-race 0.014 in	2.5 s / 453 sample	2.5 s /29 sample	2.5 s /29 sample
Inner-race 0.021 in	2.5 s / 453 sample	2.5 s /29 sample	2.5 s /29 sample
Outer-race 0.007 in	2.5 s / 453 sample	2.5 s /29 sample	2.5 s /29 sample
Outer-race 0.014 in	2.5 s / 453 sample	2.5 s /29 sample	2.5 s /29 sample
Outer-race 0.021 in	2.5 s / 453 sample	2.5 s /29 sample	2.5 s /29 sample
Ball 0.007 in	2.5 s / 453 sample	2.5 s /29 sample	2.5 s /29 sample
Ball 0.014 in	-	-	2.5 s /29 sample
Ball 0.021 in	2.5 s / 453 sample	2.5 s /29 sample	2.5 s /29 sample

rectifier activation, a batch normalization layer and a pooling layer. The first layer is composed of 24 wide filters of 128. The following convolutional layers contain each 48 filter of length seven. Pooling layers are of length three and stride three. The length of the fully connected layer is 80.Moreover, the learning rate is 1e-5, the chosen batch size is 50, and the number of epochs is 250.

This architecture is trained using three different methods. First, in the classification framework, the data are assigned to nine different classes each depicting the combined fault type and location. Second, the architecture is trained in a hierarchical way by training a model for fault isolation then training three different models each specialized in estimating the fault size for a given fault location. Finally, we train the model in a joint way as described in section 2.

The experiments were implemented using the Tensorflow deep learning framework. Furthermore, training and inference tasks are run on an Ubuntu Linux machine with a Nvidia GTX 1070 GPU.

3.3. The results

Table 2 and 3 shows the accuracy of the fault isolation and the mean absolute error of the fault size estimation on the

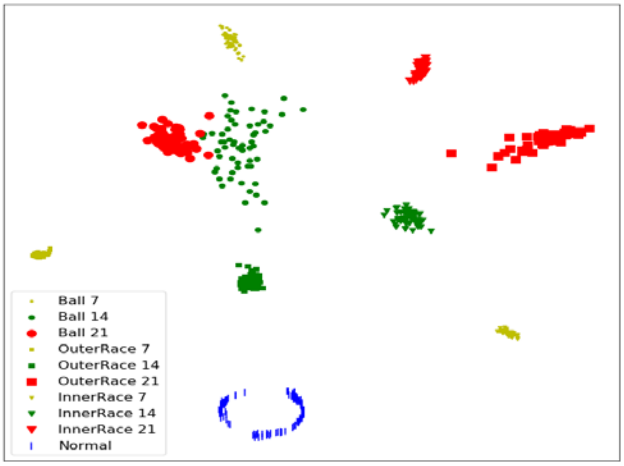

Figure 3. The Tsne representation of the classification model.

test-set respectively. Moreover, (figure 3, 4, and 5) illustrates the Tsne representation from the last fully connected layer of the classification model, the hierarchical model and the multi-task model respectively.

Table 2. The accuracy of the fault detection and isolation task.

Condition	Classification Model	Hierarchical Model	Multi-task Model
Normal	100 %	100 %	100 %
Inner-race 0.007 in	100 %	100 %	100 %
Inner-race 0.014 in	100 %	100 %	100 %
Inner-race 0.021 in	100 %	100 %	100 %
Outer-race 0.007 in	100 %	100 %	100 %
Outer-race 0.014 in	100 %	100 %	100 %
Outer-race 0.021 in	100 %	100 %	100 %
Ball 0.007 in	100 %	100 %	100 %
Ball 0.014 in	98.27 %	100 %	100 %
Ball 0.021 in	100 %	100 %	100 %

To obtain the fault size from the classification model, we select the class with the highest estimated probability. Then, we normalize the values from the selected fault type and multiply them with their nominal class. I.e. if the estimated probabilities are (0,0.1,0.9,0.05,0,0,0,0,0,0.05), the estimated fault location is the inner-race and the estimated size is $\frac{0.1}{0.1+0.9+0.05} * 0.007 + \frac{0.9}{0.1+0.9+0.05} * 0.014 + \frac{0.05}{0.1+0.9+0.05} * 0.021 = 0.01367$.

The selected neural network architecture obtains perfect in-distribution accuracy in the three different frameworks. The

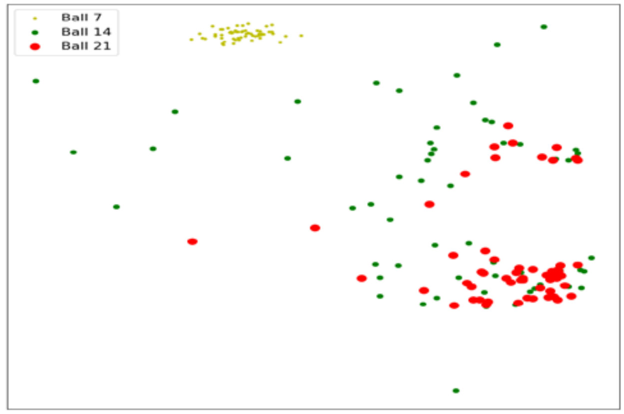

Figure 4. The Tsne representation of the hierarchical model.

accuracy here refers to the prediction of the right fault type. In the out-of-training-distribution case, the classification model performance deteriorated due to the coupling of fault size and fault type. Indeed, (figure 3) shows that the representation of each in-distribution fault-type and fault size couple are linearly separable and that the representation of the out-of-distribution fault-size is dispersed between the different clusters. This confirms the suspicion that the learned representation for the fault isolation task here is susceptible to perturbations from the fault size.

Table 3 shows the performance of the three frameworks in fault size assessment. All three approaches obtained negligible errors in the in-training-distribution cases. In contrast, large errors are observed in the case of the out-of-distribution fault size. Moreover, the multi-task approach outperforms the other frameworks significantly, as it uses information acquired from different fault types to assess the size of the unknown fault size. (Figure 5). On the other hand, the hierarchical approach model tries to extrapolate the fault size from the observed ones (figure 4) while the classification approach fails miserably.

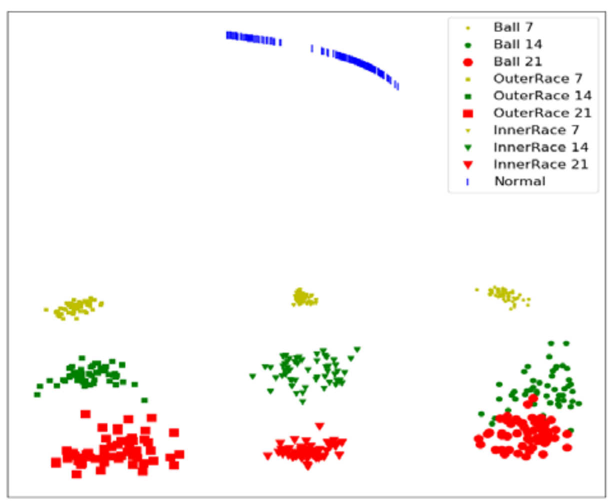

Figure 5. The Tsne representation of the multi-task model.

Table 3. The mean absolute error of fault size estimation measured in 1e-3 inch

Condition	Classification Model	Hierarchical Model	Multi-task Model
Normal	1.57e-2	2e-2	1.25e-2
Inner-race 0.007 in	1.29e-2	9e-2	0.05
Inner-race 0.014 in	3.47e-3	0.107	0.17
Inner-race 0.021 in	1.10e-2	0.113	8.5e-2
Outer-race 0.007 in	1.80e-2	0.05	0.06
Outer-race 0.014 in	8.96e-3	0.22	0.08
Outer-race 0.021 in	8.47e-2	0.12	0.25
Ball 0.007 in	2.81e-2	0.18	0.08
Ball 0.014 in	5.17	4.58	2.31
Ball 0.021 in	7.69e-2	0.55	0.17

4. CONCLUSION

In this paper, we proposed a framework for joint fault isolation and severity estimation. The presented model is based on a deep convolutional neural network and showed superior performance for out-of-distribution fault size estimation.

As the complexity of monitored assets rises, the risk of failure modes not being considered in the experimental phase increases. Therefore, the problem of open-diagnostics in its general form needs more interest from the scientific communities. Moreover, most studies do not use the data acquired during deployment to improve the performance of the intelligent diagnostics system. Our future work will focus on a lifelong learning approach for diagnostics.

REFERENCES

Ando, R.K., Zhang, T., 2005. A Framework for Learning Predictive Structures from Multiple Tasks and Unlabeled Data. Journal of Machine Learning Research 6, 1817–1853.

Chen, J., Li, Z., Pan, J., Chen, G., Zi, Y., Yuan, J., Chen, B., He, Z., 2016. Wavelet transform based on inner product in fault diagnostics of rotating machinery: A review. Mechanical Systems and Signal Processing 70–71, 1–35. https://doi.org/10.1016/j.ymssp.2015.08.023

Cohen, N., Shashua, A., 2016. Inductive Bias of Deep Convolutional Networks through Pooling Geometry. arXiv:1605.06743 [cs].

Deng, S., Cheng, Z., Li, C., Yao, X., Chen, Z., Sanchez, R.-V., 2016. Rolling bearing fault diagnostics based on Deep Boltzmann machines, in: Prognostics and System Health Management Conference (PHM-Chengdu), 2016. IEEE, pp. 1–6.

Gan, M., Wang, C., Zhu, C., 2016. Construction of hierarchical diagnostics network based on deep learning and its application in the fault pattern recognition of rolling element bearings. Mechanical Systems and Signal Processing 72–73, 92–104. https://doi.org/10.1016/j.ymssp.2015.11.014

Gebru, T., Hoffman, J., Fei-Fei, L., 2017. Fine-Grained Recognition in the Wild: A Multi-task Domain Adaptation Approach. IEEE, pp. 1358–1367. https://doi.org/10.1109/ICCV.2017.151

Guo, X., Chen, L., Shen, C., 2016. Hierarchical adaptive deep convolution neural network and its application to bearing fault diagnostics. Measurement 93, 490–502. https://doi.org/10.1016/j.measurement.2016.07.054

Guo, X., Shen, C., Chen, L., 2017. Deep Fault Recognizer: An Integrated Model to Denoise and Extract Features for Fault Diagnostics in Rotating Machinery 17.

He, J., Yang, S., Gan, C., 2017. Unsupervised Fault Diagnostics of a Gear Transmission Chain Using a Deep Belief Network 21.

Hinchi, A.Z., Tkiouat, M., 2018. Rolling element bearing remaining useful life estimation based on a convolutional long-short-term memory network. Procedia Computer Science, proceedings of the first international conference on intelligent computing in data sciences, icds2017 127, 123–132. https://doi.org/10.1016/j.procs.2018.01.106

Hochreiter, S., Schmidhuber, J., 1997. LSTM can solve hard long time lag problems, in: Advances in Neural Information Processing Systems. pp. 473–479.

Ioffe, S., Szegedy, C., 2015. Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift, in: PMLR. Presented at the International Conference on Machine Learning, pp. 448–456.

Jardine, A.K.S., Lin, D., Banjevic, D., 2006. A review on machinery diagnostics and prognostics implementing condition-based maintenance. Mechanical Systems and Signal Processing 20, 1483–1510. https://doi.org/10.1016/j.ymssp.2005.09.012

Jia, F., Lei, Y., Guo, L., Lin, J., Xing, S., 2018. A neural network constructed by deep learning technique and its application to intelligent fault diagnostics of machines. Neurocomputing 272, 619–628. https://doi.org/10.1016/j.neucom.2017.07.032

Jia, F., Lei, Y., Lin, J., Zhou, X., Lu, N., 2016. Deep neural networks: A promising tool for fault characteristic mining and intelligent diagnostics of rotating machinery with massive data. Mechanical Systems and Signal Processing 72–73, 303–315. https://doi.org/10.1016/j.ymssp.2015.10.025

LeCun, Y., Bengio, Y., Hinton, G., 2015. Deep learning. Nature 521, 436–444. https://doi.org/10.1038/nature14539

Lei, Y., Lin, J., He, Z., Zuo, M.J., 2013. A review on empirical mode decomposition in fault diagnostics of rotating machinery. Mechanical Systems and Signal Processing 35, 108–126. https://doi.org/10.1016/j.ymssp.2012.09.015

Li, S., Liu, G., Tang, X., Lu, J., Hu, J., 2017. An Ensemble Deep Convolutional Neural Network Model with Improved D-S Evidence Fusion for Bearing Fault Diagnostics. Sensors 17, 1729. https://doi.org/10.3390/s17081729

Loparo, K.A., 2005. Bearing Data Center. Case Western Reserve University.

Rai, A., Upadhyay, S.H., 2016. A review on signal processing techniques utilized in the fault diagnostics of rolling element bearings. Tribology International 96, 289–306. https://doi.org/10.1016/j.triboint.2015.12.037

Ranjan, R., Patel, V.M., Chellappa, R., 2017. HyperFace: A Deep Multi-task Learning Framework for Face Detection, Landmark Localization, Pose Estimation, and Gender Recognition. IEEE Transactions on Pattern Analysis and Machine Intelligence 1–1. https://doi.org/10.1109/TPAMI.2017.2781233

Samanta, B., Al-balushi, K.R., 2003. Artificial neural network based fault diagnostics of rolling element bearings using time-domain features. Mechanical Systems and Signal Processing 17, 317–328. https://doi.org/10.1006/mssp.2001.1462

Shao, H., Jiang, H., Zhang, H., Duan, W., Liang, T., Wu, S., 2018a. Rolling bearing fault feature learning using improved convolutional deep belief network with compressed sensing. Mechanical Systems and Signal Processing 100, 743–765. https://doi.org/10.1016/j.ymssp.2017.08.002

Shao, H., Jiang, H., Zhang, H., Liang, T., 2018b. Electric Locomotive Bearing Fault Diagnostics Using a Novel Convolutional Deep Belief Network. IEEE Transactions on Industrial Electronics 65, 2727–2736. https://doi.org/10.1109/TIE.2017.2745473

Shao, H., Jiang, H., Zhang, X., Niu, M., 2015. Rolling bearing fault diagnostics using an optimization deep belief network. Meas. Sci. Technol. 18.

Shao, H., Jiang, H., Zhao, H., Wang, F., 2017. A novel deep autoencoder feature learning method for rotating machinery fault diagnostics. Mechanical Systems and Signal Processing 95, 187–204. https://doi.org/10.1016/j.ymssp.2017.03.034

Tao, J., Liu, Y., Yang, D., 2016. Bearing Fault Diagnostics Based on Deep Belief Network and Multisensor Information Fusion. Shock and Vibration 2016, 1–9. https://doi.org/10.1155/2016/9306205

Wang, Y., Xiang, J., Markert, R., Liang, M., 2016. Spectral kurtosis for fault detection, diagnostics and prognostics of rotating machines: A review with applications. Mechanical Systems and Signal Processing 66–67, 679–698. https://doi.org/10.1016/j.ymssp.2015.04.039

Yan, R., Gao, R.X., Chen, X., 2014. Wavelets for fault diagnostics of rotary machines: A review with applications. Signal Processing, Time-frequency methods for condition based maintenance and modal analysis 96, 1–15. https://doi.org/10.1016/j.sigpro.2013.04.015

Yang, J., Zhang, Y., Zhu, Y., 2007. Intelligent fault diagnostics of rolling element bearing based on SVMs and fractal dimension. Mechanical Systems and Signal Processing 21, 2012–2024. https://doi.org/10.1016/j.ymssp.2006.10.005

Young, S.R., Rose, D.C., Karnowski, T.P., Lim, S.-H., Patton, R.M., 2015. Optimizing Deep Learning Hyper-parameters Through an Evolutionary Algorithm, in: Proceedings of the Workshop on Machine Learning in High-Performance Computing Environments, MLHPC ’15. ACM, New York, NY, USA, pp. 4:1–4:5. https://doi.org/10.1145/2834892.2834896

Zhang, W., Li, C., Peng, G., Chen, Y., Zhang, Z., 2018. A deep convolutional neural network with new training methods for bearing fault diagnostics under noisy environment and different working load. Mechanical Systems and Signal Processing 100, 439–453. https://doi.org/10.1016/j.ymssp.2017.06.022

Zhang, W., Peng, G., Li, C., Chen, Y., Zhang, Z., 2017. A New Deep Learning Model for Fault Diagnostics with Good Anti-Noise and Domain Adaptation Ability on Raw Vibration Signals. Sensors 17, 425. https://doi.org/10.3390/s17020425

Zoph, B., Le, Q.V., 2016. Neural Architecture Search with Reinforcement Learning. arXiv:1611.01578 [cs].

BIOGRAPHIES

Ahmed Zakariae Hinchi is currently a Ph.D. student in the laboratory of applied mathematics (LERMA) in the Mohammadia School of engineering. He holds an engineering degree in Industrial engineering and integrated manufacturing from the ENSAM School of engineering. His research interests include prognostics and health management, reliability, machine learning, and deep learning.

Mohamed Tkiouat is a Professor at the Mohammadia School of Engineering, Mohamad V University of Rabat, Morocco. His primary areas of research include Markovian and multi-agent models, reliability and risk management.