ORIGINAL ARTICLE

A novel deep learning approach for intelligent fault diagnosis applications based on time-frequency images

Özgür Gültekin[1,3] · Eyüp Çinar[2,3] · Kemal Özkan[2,3] · Ahmet Yazıcı[2,3]

Abstract

Deep learning (DL) techniques have been gaining ground for intelligent equipment/process fault diagnosis applications. However, employing DL methods for such applications comes with its technical challenges. The DL methods are utilized to extract features from raw data automatically, which leads up to its own complications in data preprocessing and/or feature engineering phases. Moreover, another difficulty arises when DL methods are employed utilizing single type of sensor data as the performance of a fault diagnosis application is hindered. To address these issues, we propose utilization of a deep residual network-based multi-sensory data fusion method. The method is established on time-frequency images obtained by short-time Fourier transform to diagnose machine faults. The experimental results demonstrate that the proposed model combining different types of measured signals can diagnose bearing conditions on machines more effectively compared to a single type of measured signal in terms of diagnostic accuracy.

Keywords Bearing fault diagnosis · Deep learning · Deep residual network · Predictive maintenance · Short-time Fourier transform

1 Introduction

Intelligent fault diagnosis applications have gained considerable attraction in manufacturing. These applications are implemented via computational machine learning systems and data processing algorithms that can detect anomalies in industrial machines and accurately predict faults before they occur. In this way, the risk of failure can be predicted in advance for industrial automation systems. Diagnosed faults can be integrated into manufacturers' predictive maintenance solutions to increase production quality and capacity.

✉ Eyüp Çinar
  eyup.cinar@ogu.edu.tr

1 Department of Informatics, Eskişehir Osmangazi University, Meşelik Campus, 26040 Eskişehir, Turkey

2 Department of Computer Engineering, Eskişehir Osmangazi University, Meşelik Campus, 26040 Eskişehir, Turkey

3 Center of Intelligent Systems Application Research, Eskisehir Osmangazi University, Meşelik Campus, 26040 Eskişehir, Turkey

Data-driven methods incorporating machine learning (ML) models are at the heart of these intelligent fault diagnosis applications [1–3]. Having emerged as a subclass of ML models in recent years, deep learning (DL) models show promising results on intelligent fault detection applications for machine parts [4]. This is primarily due to their power features with robustness against the noise in the data and their ability to learn features from measurement data automatically. It bears great potential to help ease complex feature extraction, elimination, and select procedures needed while processing a large amount of historical time-series sensor data [5].

The DL approaches for intelligent fault diagnosis applications often utilize a single type of sensor signal collected over a long historical time frame. A large volume of historical data collected from a sensor is supplied to a DL model. The fault detection and classification tasks are performed on unobserved data with the incorporated method. Promising results have been obtained utilizing DL techniques; however, it is often times the case that it is not viable to obtain such a large volume of historical data all the time in a manufacturing environment due to various changes in equipment baseline or process drifts. Therefore,

developing a robust multi-sensory fusion technique is quite advantageous where multiple sensor measurements collected from different sources on a machine can be processed and utilized in a single model [6]. In this work, we propose a novel DL method for a fault diagnosis application where multiple data sources can be monitored simultaneously. The main contributions of this study can be listed as:

- A novel deep residual network (DRN) based on multi-model data fusion method for intelligent fault diagnosis is proposed.
- Time–frequency representations of simultaneously measured non-stationary signals acquired by short-time Fourier transform (STFT) are fed into the separate identical DRN models and fused.
- The proposed method is tested on challenging datasets to observe its robustness.
- The experimental results indicate that the proposed fusion method achieves higher performance whenever the dataset is getting complex and bigger.

The paper is organized as follows: Firstly, a literature review on related work for DL approaches targeting similar fault diagnosis applications is introduced. Secondly, the details of the proposed method and model design process are presented. Thirdly, the experimental work and the results obtained by the novel approach are presented with a discussion. Lastly, the conclusions and future work that will follow this study are given.

2 Background and literature review

One of the frequently used DL models in intelligent fault diagnosis is convolutional neural network (CNN) architecture which employs multiple processing layers to process data in multiple arrays and extracts features. DRN, a variant of CNN, is also used by many researchers to diagnose rotating machinery faults. Since DRNs utilize identity shortcuts, it is easier to optimize parameters and decrease the possibility of overfitting in deeper models. Zhang et al. [7] employed raw vibration signals as the input of their proposed DRN-based method to diagnose bearing faults. Results of the study indicate that the proposed method obtained higher testing diagnosis accuracy than other traditional CNN-based methods that are cross compared in the literature. In [8], Ma et al. used time–frequency representations and DRNs in their proposed data-driven fault diagnosis method which is applied to planetary gearboxes. The results show that their proposed method outperformed the other methods compared in terms of diagnostic accuracy. Zhao et al. [9] designed a variant of DRN that utilized dynamically weighted wavelet

coefficients to diagnose faults of planetary gearboxes. Their proposed method achieves better training and testing accuracy (in both noisy and noiseless environments) as opposed to the methods incorporating shallow and/or deep learning algorithms.

Many researchers employ signal processing techniques to transform a measured signal between different domains like time, frequency, or time–frequency to utilize time-series sensor data better. Even if there is no exact answer to which domain would be more suitable for intelligent diagnosis, many studies show that time–frequency representations have better performance since they include much more information than only time or frequency-domain representations. For example, Pandhare et al. [10] used time, frequency, and time–frequency domain representations as inputs to train the CNN model to diagnose bearing faults. Their experiment results have shown that the model fed by time–frequency inputs obtained by STFT had better accuracy than other models. Wang et al. [11] preprocessed raw vibration signals by STFT to acquire corresponding time–frequency maps and fed their CNN model to diagnose motor faults. Results of the work demonstrate that CNN with time–frequency map inputs show better performance than other methods compared in the literature in terms of diagnostic accuracy. In [12], Zhang et al. used STFT to obtain input images of their proposed LeNet-5 CNN model to diagnose bearing faults. Their proposed method achieves better training and testing accuracy compared to both time domain and frequency domain methods.

The multi-sensor data fusion in an intelligent diagnosis application is considered a promising technique since the information obtained from multi-sensors can be much more meaningful than a single source [13]. Therefore, using more than one measured signal helps develop a more effective and robust fault diagnosis approach. For example, Jing et al. [14] constructed an adaptive multi-sensor data fusion method based on deep CNN (DCNN) for intelligent planetary gearbox fault detection, which takes raw signal data as input. In this work, they used four types of measured signals like vibration, acoustics, current, and instantaneous angular speed. According to the results, their proposed model that employs data fusion achieved better testing accuracy than the other methods evaluated in their experiments. Wei et al. [15] have recently proposed a data fusion method that takes advantage of data measured from multi-sensors to detect incipient faults. The experiments performed on aircraft engines show that their proposed data fusion method can detect occurrences of incipient faults more accurately and robustly. Xu et al. [16] developed a novel integrated model named parallel CNN (PCNN), which benefits from multi-sensory feature fusion and popular DL algorithms to monitor cutting tools and

diagnose bearing faults. Their results on two different experiments indicate that the proposed approach provides more accuracy and effectiveness on both pattern classification and regression prediction problems. Even if signals measured from multiple sensors have more meaning than single-sensor signals, most of the literature studies cannot use multiple sensors effectively. Some of the studies focused on only one type of signal obtained by multiple sensors or combined different types of raw signals before feeding their models. Alternatively, we propose a novel DRN-based, multi-sensory data fusion method, which takes time–frequency representations generated by means of gathering different types of measured signals as input, and fuses relevant features extracted prior to classification process.

3 Proposed method

This section describes the components and structure of our proposed method. We first briefly explain the multi-sensory data fusion approach. Then we provide an overview of deep residual learning, which constitutes the building blocks of our convolutional neural network structure. Finally, we introduce the proposed DRN-based multi-model data fusion method.

3.1 Multi-sensory data fusion

A system is a combination of many components that work in harmony. Each component has different characteristics and status that represent overall system conditions. Therefore, analyzing system conditions based on only a single type of measured signal will potentially lead to inadequate analyses in certain situations, where near miss or miss occurrences are present. To address this issue, one can expand the diagnosis step and start to monitor multiple types of sensors. Fusing information gathered by multiple measurement sources concurrently might help monitor a system better and increase the fault diagnosis model capacity [17]. Since different sensors have different advantages (or drawbacks), a multi-sensor fusion approach can strengthen the evidence on failure modes as opposed to single-sensor monitoring cases.

3.2 Deep residual learning

Deep residual learning is a learning process that benefits from residual networks. The residual networks consist of residual learning blocks and shortcut connections. Each residual learning block uses some weight layers, batch normalization (BN), and nonlinearity functions to learn from inputs. Moreover, each shortcut connection is used to skip some of the layers and add input to the output of the learning block element-wise before the last nonlinearity function. Figure 1 demonstrates a residual building block of a DRN model, consisting of double weight layers, BN, and ReLU nonlinearity function.

As shown in the figure, the input of the first weight layer is added to the output of the last layer before ReLU function is applied. In addition, if input and output dimensions are different, a projection operation should be inserted to match them at the identity shortcut connection. This approach helps the model recall inputs and solves the degradation problem of deep networks [18].

3.3 Model design

In the field of intelligent fault diagnosis, most of the studies have been focusing on a single type of measured signal such as vibration, temperature, or current alone. However, each different type of measured signal might carry fingerprints that present discriminative information about machinery parts being monitored. In this paper, we propose a multi-model structure that is able to use more than one type of measured signal simultaneously as separate inputs for separate DRNs and then fuse them before classification. In an effort to accomplish our goal, we utilize DRN structures as a building block [18]. As shown in Fig. 2, the proposed multi-model approach consists of multiple ResNet-18 models. Each model takes time–frequency representations of a single type of measured signal as an input and then all of the features generated by multiple models are combined before classification. Each ResNet-18 model has 18 layers. The first layer is 3×3 convolutions
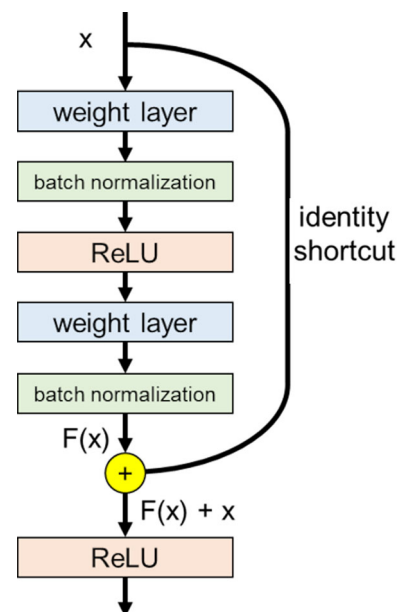

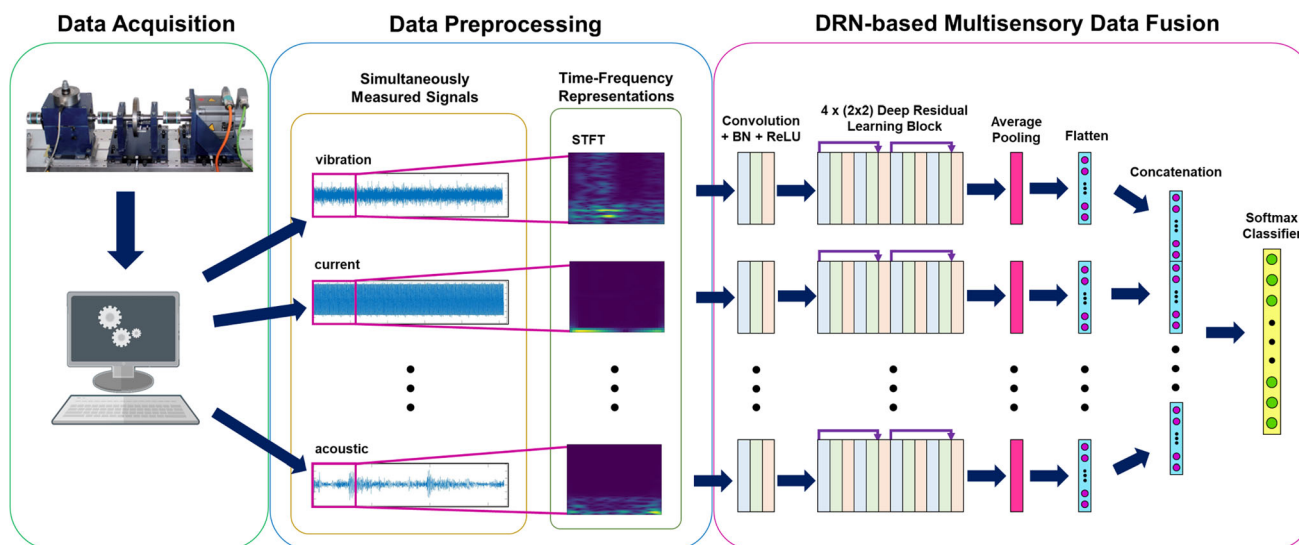**Fig. 1** Structure of a residual learning block

Fig. 2 Structure of the proposed method

with a stride of 1 and 16 filters. Then, there is a stack of 4 layers with 3 × 3 convolutions 4 times. Each of the layers in a stack consists of a convolution, BN, and activation steps. At the end of each stack, the size of the feature maps is halved, and the number of filters is doubled. The down-sampling is done by convolutions with a stride of 2. After those layers, a global average pooling with the size of 8 is applied for flattening. As the last step, outputs of each model are combined with concatenation operation and connected to the same classification layer for the final prediction. All activation functions used in residual learning blocks are ReLU function and Softmax function is used as the classification function at the output layer.

Figure 3 shows a flowchart of the proposed method.

4 Experimental results and analysis

4.1 Dataset descriptions

A publicly available dataset called Paderborn University (PU) bearing dataset [19, 20], which is provided by KAt-DataCenter of Paderborn University, is used in this study. The modular test rig of the PU bearing dataset is presented in Fig. 4. The dataset has 32 sets of motor current and vibration signals. Each set has different aspects; (a) six of them are measured from healthy bearings, (b) 12 of them are measured from artificially damaged bearings and (c) 14 of them are measured from actually damaged bearings. In addition to this, the test rig is operated under four different operating conditions; (1) runs at 1500 rpm with a load torque of 0.7 Nm and a radial force on the bearing of 1000 N, (2) runs at 900 rpm with a load torque of 0.7 Nm and a radial force on the bearing of 1000 N, (3) runs at

1500 rpm with a load torque of 0.1 Nm and a radial force on the bearing of 1000 N and (4) runs at 1500 rpm with a load torque of 0.7 Nm and a radial force on the bearing of 400 N. Each set has 20 measurements of 4 s for each aforementioned operating condition. Therefore, each of the 32 sets has 80 different files for four conditions.

In this work, four datasets are organized and used to test the proposed method's efficiency, effectiveness, and performance. The datasets are carefully chosen in accordance with the number of bearings from which the signals are collected with specific conditions. Since the proposed method can learn effectively from multiple signals, we aim to benchmark it on relatively more challenging tasks. Therefore, dataset A, B, and C are composed of measured signals obtained from similarly conditioned bearings as six healthy, 12 artificially damaged, and 14 actually damaged. Moreover, dataset D also comprises measured signals from 32 bearings as a combination of six healthy, 12 artificially damaged, and 14 actually damaged ones. Besides, to test the method's performance on different data sizes, the number of bearing conditions in each dataset is specified unequally. Since each dataset contains a different number of samples and conditions, it is possible to evaluate the performance of the proposed method on both small and large datasets.

Additionally, each dataset includes measured signals from the setting that runs at 1500 rpm with a load torque of 0.7 Nm and a radial force of 1000 N on the bearing. Dataset A includes 4 s of measured signal from 6 of the healthy bearings. Dataset B contains 4 s of measured signals from 12 artificially damaged bearings. Dataset C includes 4 s of measured signals from 14 actually damaged bearings. Dataset D contains 4 s of measured signals from 32 different bearings. Each dataset mentioned above has
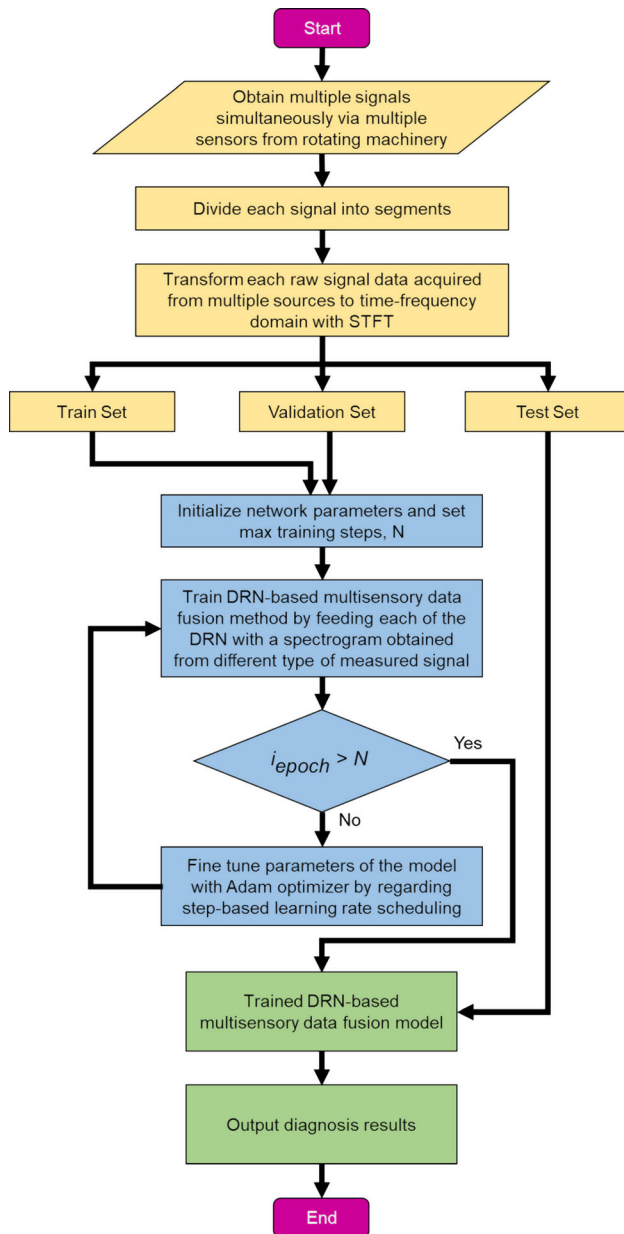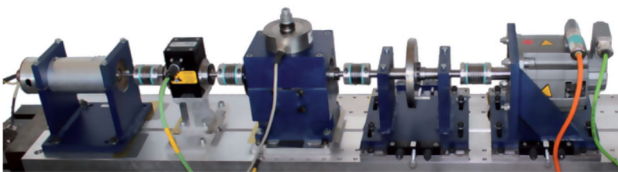
Fig. 3 Flowchart of the proposed method

Fig. 4 Modular test rig of PU bearing dataset

800 samples (400 for current signals and 400 for vibration signals) with a length of 512 for each bearing included. Therefore, dataset A contains 4800, dataset B contains 9600, dataset C contains 11200, and dataset D contains

25600 samples in total. Detailed information about datasets is given in Table 1.

4.2 Experimental setting

The four bearing datasets described in the previous section include both current and vibration signals, which are used to test the performance of the proposed model. Since input image size selection is a critical process of any CNN model, we used these parameters and tested them, which resulted in high accuracy and optimal time complexity as defined in [12, 21]. However, the number of overlapping points is excluded such that time–frequency representations for STFT are obtained.

In the preparation stage, each measured signal from bearings was divided into samples with a length of 512, measured in 8 ms. A sliding window accomplished the dividing process, and there were no overlapping points. After taking 400 samples from each measured signal, all of the samples were converted to the time–frequency resolutions with the parameters given in Table 2. By such parameters, each 512 points was converted to a 65×65 spectrogram. To make the convolutions easier with the proposed model, we truncated each spectrogram's last row and column. Therefore, this operation determined our input image size as 64×64, and there were 400 of them for each measured signal. After the preparation stage, each bearing dataset was randomly divided into training, validation, and test sets with the ratio of 0.6, 0.2, and 0.2, respectively. In deep models, training and validation sets were used for the learning and test sets for the prediction phase. Considering time consumption and overtraining issues, we specified the size of each batch as 32 and the number of the epoch as 200. Moreover, the Adam stochastic optimization algorithm [22] was used with step-based learning rate scheduling that decreased after every few epochs [23]. In this experiment, for the step-based learning rate function, the parameters initial learning rate, base drop rate, and step size were specified as 0.01, 0.1, and 40, respectively. In addition, the learning rate is specified as 0.001 for the first 40 epochs.

The code is written with Python 3.6.9 and TensorFlow (version 2.3.0) library is used to build deep models. The computer also has a Tesla T4 14.64 GB GPU and 12.72 GB RAM.

4.3 Classification performance

The experiments performed in this study consist of three sequential steps. In the first two steps, we performed our diagnosis model for only a single type of measured signal, and as the last step, we applied the proposed multi-model approach to two types of different signals. Then, we

Table 1 Description of datasets

Dataset	Data description	Bearing codes	Data type	Number of STFT samples for each bearing code (current/vibration)
A	6 healthy bearings	K001, K002, K003, K004, K005, K006	Current, vibration	400/400 (2400/2400 in total)
B	12 artificially damaged bearings	KA01, KA03, KA05, KA06, KA07, KA08, KA09, KI01, KI03, KI05, KI07, KI08	Current, vibration	400/400 (4800/4800 in total)
C	14 actually damaged bearings	KA04, KA15, KA16, KA22, KA30, KB23, KB24, KB27, KI04, KI14, KI16, KI17, KI18, KI21	Current, vibration	400/400 (5600/5600 in total)
D	32 bearings as combination of healthy, artificially damaged and actually damaged ones	All of the bearing codes above	Current, vibration	400/400 (12800/12800 in total)

Table 2 Selected parameters for STFT

Window length	Overlapping points	FFT points	Time–frequency resolution	Input image size
128	120	128	65×65	**64×64**

compared obtained results in terms of testing accuracy and classification performance. To test the robustness and generalization capability of the proposed method, we performed each step for 15 times on each dataset and took the average of the results as final results. The diagnosis results of these three steps are given in Fig. 5. Furthermore, the proposed DRN structure details are also given in Table 3 for a better understanding of the proposed method.

In the first step, we implemented our model for only vibration signals from dataset A-D to diagnose bearings' conditions. Diagnostic accuracies obtained from 15 trials were greater than 85.62%, 93.65%, 96.52%, and 90.98% for dataset A, B, C, and D, respectively. Also, maximum accuracies acquired from the vibration signals were 89.58%, 95.83%, 98.12%, and 92.77% for dataset A, B, C, and D, respectively. In addition, the calculated standard
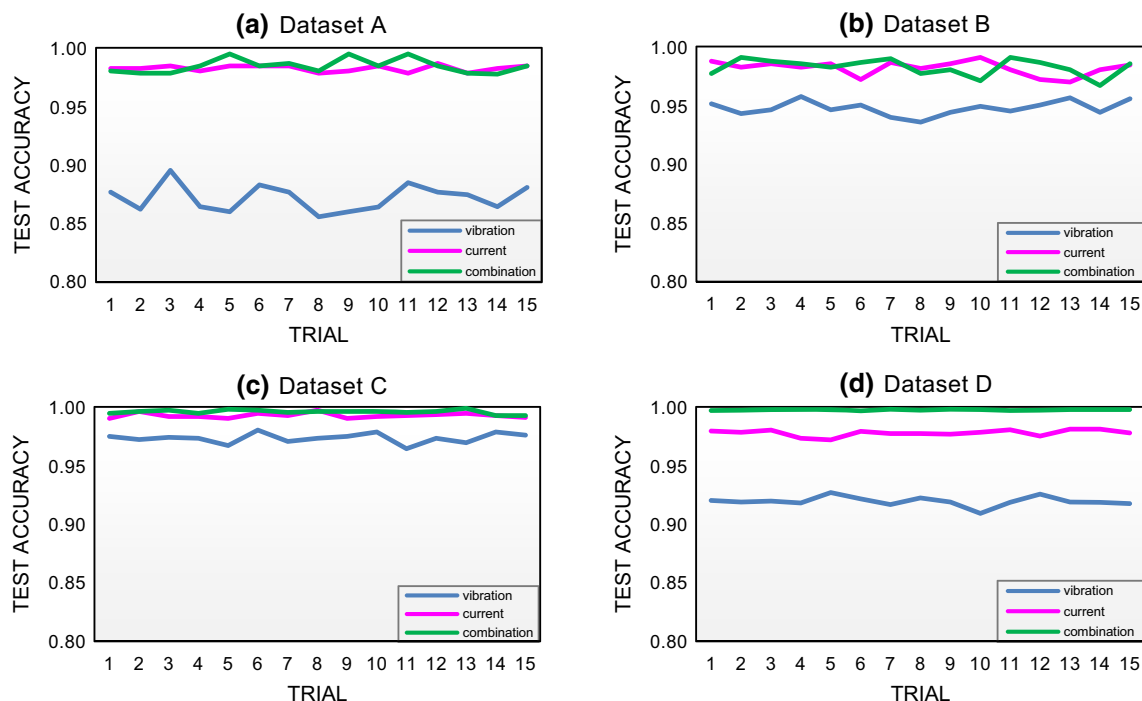

Fig. 5 Diagnosis results of 15 trials of bearing datasets using the proposed method: (**a**) Dataset A, (**b**) Dataset B, (**c**) Dataset C and (**d**) Dataset D

Table 3 Structure of deep residual networks

Layer type	Kernel size	Stride	Channel size	Feature map size	Number of models
Input	–	–	–	64×64	m
Convolution1	3×3	1	16	$64 \times 64 \times 16$	m
Convolution2	$\begin{bmatrix} 3 \times 3 \\ 3 \times 3 \end{bmatrix} \times 2$	1, 1, 1, 1	16	$64 \times 64 \times 16$	m
Convolution3	$\begin{bmatrix} 3 \times 3 \\ 3 \times 3 \end{bmatrix} \times 2$	2, 1, 1, 1	32	$32 \times 32 \times 32$	m
Convolution4	$\begin{bmatrix} 3 \times 3 \\ 3 \times 3 \end{bmatrix} \times 2$	2, 1, 1, 1	64	$16 \times 16 \times 64$	m
Convolution5	$\begin{bmatrix} 3 \times 3 \\ 3 \times 3 \end{bmatrix} \times 2$	2, 1, 1, 1	128	$8 \times 8 \times 128$	m
Global Pooling	8×8	1	128	$1 \times 1 \times 128$	m
Flatten	–	–	–	128	m
Concatenate	–	–	–	128*m	1
Softmax	–	–	–	n	1

m denotes number of models and n denotes number of classes

deviations were 1.14%, 0.62%, 0.43% and 0.41% for dataset A, B, C and D, respectively.

In the second step, we selected only samples generated from current signals in each dataset. When we applied our DRN model to the current signals in each dataset 15 times, our diagnostic accuracies were greater than 97.92%, 97.08%, 99.11%, and 97.27% for dataset A, B, C, and D, respectively. Maximum accuracies for the current signal samples were 98.75%, 99.17%, 99.82%, and 98.20% for dataset A, B, C, and D, respectively. In addition, standard deviations of 15 trials were 0.27%, 0.61%, 0.22%, and 0.28% for dataset A, B, C, and D, respectively.

In the last step of our experiment, the proposed multi-model method was applied to each dataset 15 times. Samples consisting of both vibration, and current signals were used as inputs. All of the diagnostic accuracies acquired were greater than 97.83%, 96.77%, 99.37%, and 99.77% for dataset A, B, C, and D, respectively. Moreover, maximum accuracies reached were 99.58%, 99.17%, 100%, and 99.92% for datasets A, B, C, and D, respectively. In addition, standard deviations of 15 trials were 0.62%, 0.71%, 0.18%, and 0.04% for dataset A, B, C, and D, respectively. These results show that our model can achieve high fault diagnosis accuracy rate on test set, suggesting that the model can learn more information when both vibration and current signals are simultaneously utilized along with time–frequency representations. The average accuracies and standard deviations of 15 trials are presented in Table 4 to help the reader compare the results of these three processes. The average accuracies for different datasets demonstrate that the proposed multi-model approach can diagnose bearing conditions with high accuracy by benefitting from more than one type of measured signal. In addition, low standard deviations manifest the robustness of the proposed model as well.

Figure 6 demonstrates the diagnosis results on dataset A-D in one of the 15 trials using confusion matrices. Each confusion matrix manifests classification results obtained on samples, including both vibration and current signals. Labels on each matrix indicate the code of bearings. Evaluating the scores presented via confusion matrices also shows us that the proposed multi-model method can achieve the diagnosis of each bearing condition with high accuracy even if the number of bearings is increased. However, we did not observe the same trend when a single type of measured signal is used, indicating that a higher accuracy can be obtained with the proposed multi-sensor fusion approach.

4.4 Discussion

A method with the advantage of incorporating more than one type of measured signal is designed in this study. The method is benefitting from the STFT, the DRN, and data fusion operations. The STFT provides both time and frequency information of the measured signals. The DRN extract features from the time–frequency representations and solves the degradation problem by recalling inputs. The data fusion operation increases the number of utilized features to diagnose machine conditions. Each sensor has certain advantages and can reflect the machine conditions from different perspectives. Therefore, analyzing only one type of measured signal is insufficient to detect faults in some cases. The results obtained from the experiments demonstrate that it is possible to overcome the mentioned problem by the fusion of different types of measured signals. The datasets A-C were organized such that each group

Table 4 Comparison of average accuracy and standard deviation

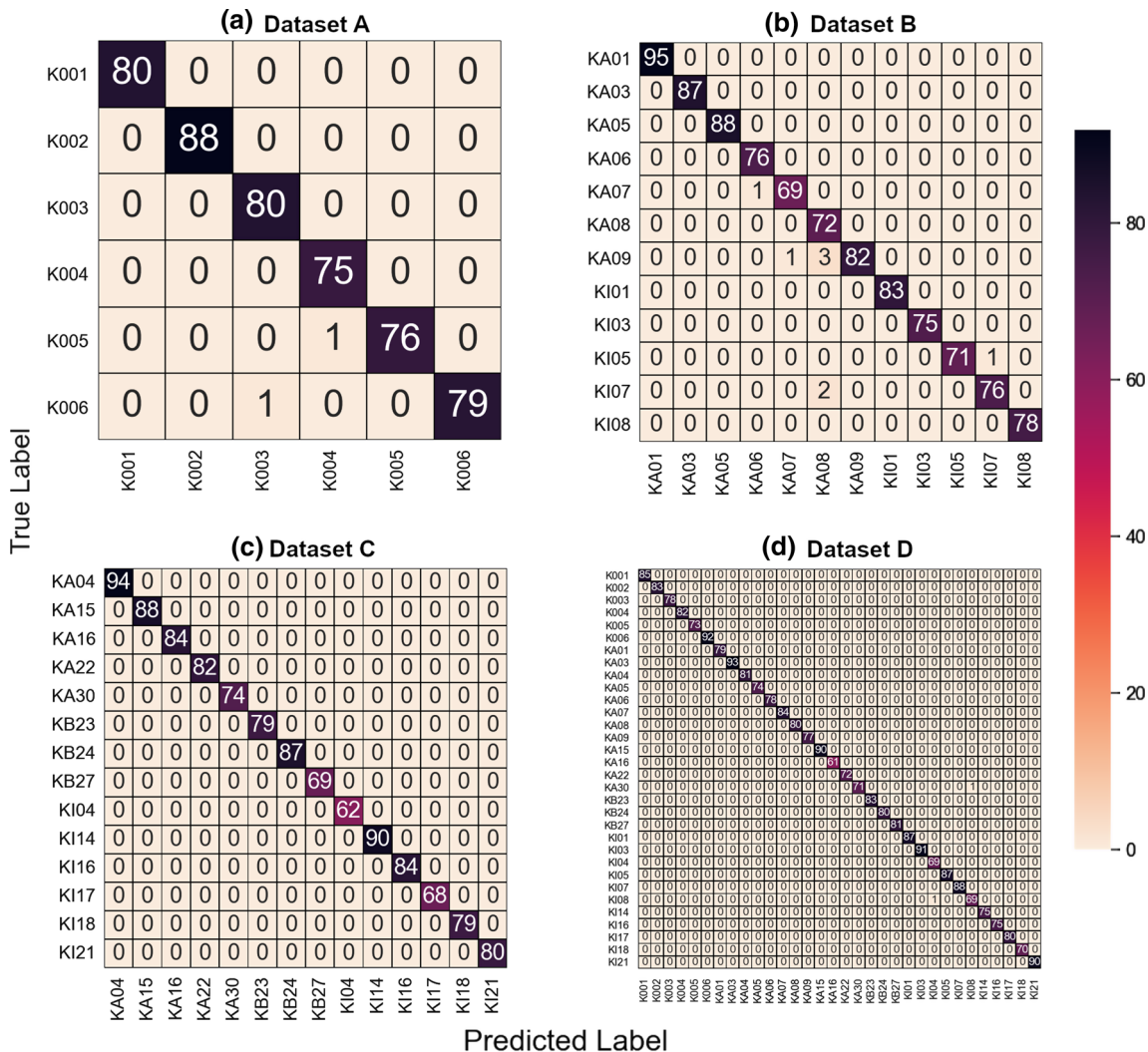| Dataset	Only vibration signals		Only current signals		Proposed method (vibration and current)	
	Testing accuracy (%)	Standard deviation (%)	Testing accuracy (%)	Standard deviation (%)	Testing accuracy (%)	Standard deviation (%)
A	87.24	1.14	98.33	0.27	98.53	0.62
B	94.84	0.62	98.27	0.61	98.35	0.71
C	97.43	0.43	99.38	0.22	99.69	0.18
D	92.01	0.41	97.88	0.28	99.86	0.04

Fig. 6 Confusion matrices generated by applying the proposed method on each dataset: (**a**) Dataset A, (**b**) Dataset B, (**c**) Dataset C and (**d**) Dataset D

consists of the data from the similar bearing statuses and characteristics such as all healthy, all artificially damaged, or actually damaged. Therefore, diagnosing the condition of such bearings is not straightforward. Although relative testing accuracies were obtained for these challenging

datasets at each of the compared steps, the method provided higher testing accuracies in most of the trials utilizing multiple signals. Whenever the dataset is getting more complex and larger like dataset D, the methods' effectiveness can be observed clearly. Our results indicate

that the model may learn better via the combination of features extracted from different types of measured signals and diagnose faults more accurately and robustly. Additionally, as the number of bearing fault modes increases and the dataset size gets larger, we observe that the proposed method can still diagnose the faults with higher accuracies as opposed to when only a single type of measured signal is utilized.

5 Conclusions and future work

This paper presents a novel DRN-based multi-model approach using time–frequency representations obtained by STFT as inputs to diagnose faults among varying bearing conditions. To verify the effectiveness and robustness of the proposed method, four datasets provided by PU bearing conditions were used in the experiments. Multiple experiments have been performed to test the robustness of the proposed method. The results demonstrate that the proposed multi-model approach can produce high diagnostic accuracies for the bearing conditions and the average testing accuracies are more than 98.35%, with considerably low standard deviations for challenging datasets. Moreover, when the number of fault conditions increased, the testing accuracy acquired by a single type of measured signal decreased dramatically, especially for 32 classes case. As has been demonstrated, the proposed method, which utilizes more than one type of measured signal, is significantly robust and can diagnose bearing conditions effectively even if the diagnosing tasks are complex and challenging. The results mentioned above justify that the method has a great potential for intelligent fault diagnosis. In addition, the flexible and adaptable nature of the method provides the opportunity for fault diagnosis of various machines utilized in industrial environments.

In future work, the proposed method will be adapted to other fault scenarios of different device components to test the validation of the intelligent diagnosis approach. In addition, similar to the testbed utilized in this study, a new testbed is designed and manufactured to collect data from new types of multiple signals. Similar experiments will be carried out with new sensors and the data to evaluate the proposed method for different fault types. Furthermore, performing diagnoses on real-time data acquired from industrial machines via the proposed approach will also be focused on.

Author's contributions Conceptualization: Eyüp Çinar and Kemal Özkan; Methodology: Kemal Özkan; Formal analysis and investigation: Özgür Gültekin and Eyüp Çinar; Writing—original draft preparation: Özgür Gültekin; Writing—review and editing: Eyüp Çinar and Kemal Özkan; Funding acquisition: Eyüp Çinar; Software: Özgür Gültekin; Resources: Ahmet Yazıcı; Supervision: Ahmet Yazıcı, Project administration: Eyüp Çinar.

Availability of data and material The data that supports the findings of this study is available at https://mb.uni-paderborn.de/en/kat/main-research/datacenter/bearing-datacenter/data-sets-and-download (accessed on November 2020). The source code that is typed for the proposed method and used in the experiments is available at https://github.com/ozggultekin/MultisensoryDataFusionWithSTFT.

Declarations

Conflict of interest The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

1. Carvalho TP, Soares FA, Vita R, Francisco RdP, Basto JP, Alcalá SG (2019) A systematic literature review of machine learning methods applied to predictive maintenance. Comput Ind Eng 137:106024. https://doi.org/10.1016/j.cie.2019.106024
2. Kumar P, Hati AS (2020) Review on machine learning algorithm based fault detection in induction motors. Arch Comput Methods Eng. https://doi.org/10.1007/s11831-020-09446-w
3. Lei Y, Yang B, Jiang X, Jia F, Li N, Nandi AK (2020) Applications of machine learning to machine fault diagnosis: a review and roadmap. Mech Syst Signal Process 138:106587. https://doi.org/10.1016/j.ymssp.2019.106587
4. Hoang D-T, Kang H-J (2019) A survey on deep learning based bearing fault diagnosis. Neurocomputing 335:327–335. https://doi.org/10.1016/j.neucom.2018.06.078
5. Hinton GE, Salakhutdinov RR (2006) Reducing the dimensionality of data with neural networks. Science 313(5786):504–507. https://doi.org/10.1126/science.1127647
6. Zhang W, Yang D, Wang H (2019) Data-driven methods for predictive maintenance of industrial equipment: a survey. IEEE Syst J 13(3):2213–2227. https://doi.org/10.1109/JSYST.2019.2905565
7. Zhang W, Li X, Ding Q (2019) Deep residual learning-based fault diagnosis method for rotating machinery. ISA Trans 95:295–305. https://doi.org/10.1016/j.isatra.2018.12.025
8. Ma S, Chu F, Han Q (2019) Deep residual learning with demodulated time-frequency features for fault diagnosis of planetary gearbox under nonstationary running conditions. Mech Syst Signal Process 127:190–201. https://doi.org/10.1016/j.ymssp.2019.02.055
9. Zhao M, Kang M, Tang B, Pecht M (2017) Deep residual networks with dynamically weighted wavelet coefficients for fault diagnosis of planetary gearboxes. IEEE Trans Industr Electron 65(5):4290–4300. https://doi.org/10.1109/TIE.2017.2762639
10. Pandhare V, Singh J, Lee J Convolutional neural network based rolling-element bearing fault diagnosis for naturally occurring and progressing defects using time-frequency domain features. In: 2019 Prognostics and System Health Management Conference (PHM-Paris), 2019. IEEE, pp. 320–326. Doi: https://doi.org/10.1109/PHM-Paris.2019.00061

11. Wang L-H, Zhao X-P, Wu J-X, Xie Y-Y, Zhang Y-H (2017) Motor fault diagnosis based on short-time Fourier transform and convolutional neural network. Chin J Mech Eng 30(6):1357–1368. https://doi.org/10.1007/s10033-017-0190-5

12. Zhang Y, Xing K, Bai R, Sun D, Meng Z (2020) An enhanced convolutional neural network for bearing fault diagnosis based on time–frequency image. Measurement. https://doi.org/10.1016/j.measurement.2020.107667

13. Qiao H, Wang T, Wang P, Qiao S, Zhang L (2018) A time-distributed spatiotemporal feature learning method for machine health monitoring with multi-sensor time series. Sensors 18(9):2932. https://doi.org/10.3390/s18092932

14. Jing L, Wang T, Zhao M, Wang P (2017) An adaptive multi-sensor data fusion method based on deep convolutional neural networks for fault diagnosis of planetary gearbox. Sensors 17(2):414. https://doi.org/10.3390/s17020414

15. Wei Y, Wu D, Terpenny J (2020) Robust incipient fault detection of complex systems using data fusion. IEEE Trans Instrum Meas 69(12):9526–9534. https://doi.org/10.1109/TIM.2020.3003359

16. Xu X, Tao Z, Ming W, An Q, Chen M (2020) Intelligent monitoring and diagnostics using a novel integrated model based on deep learning and multi-sensor feature fusion. Measurement. https://doi.org/10.1016/j.measurement.2020.108086

17. Varshney PK (1997) Multisensor data fusion. Electron Commun Eng J 9(6):245–253. https://doi.org/10.1049/ecej:19970602

18. He K, Zhang X, Ren S, Sun J (2016) Deep residual learning for image recognition. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 770–778. Doi:https://doi.org/10.1109/CVPR.2016.90

19. Lessmeier C, Kimotho J, Zimmer D, Sextro W (2019) KAt-DataCenter, chair of design and drive technology. Paderborn University, Paderborn

20. Lessmeier C, Kimotho JK, Zimmer D, Sextro W (2016) Condition monitoring of bearing damage in electromechanical drive systems by using motor current signals of electric motors: a benchmark data set for data-driven classification. In: Proceedings of the European conference of the prognostics and health management society, pp. 05–08

21. Meng Z, Guo X, Pan Z, Sun D, Liu S (2019) Data segmentation and augmentation methods based on raw data using deep neural networks approach for rotating machinery fault diagnosis. IEEE Access 7:79510–79522. https://doi.org/10.1109/ACCESS.2019.2923417

22. Kingma DP, Ba J (2014) Adam: a method for stochastic optimization. arXiv preprint arXiv: 1412.6980

23. Zagoruyko S, Komodakis N (2016) Wide residual networks. arXiv preprint arXiv: 1605.07146

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.