Frontiers of Information Technology & Electronic Engineering
www.jzus.zju.edu.cn; engineering.cae.cn; www.springerlink.com
ISSN 2095-9184 (print); ISSN 2095-9230 (online)
E-mail: jzus@zju.edu.cn

Supplementary materials for

Binkun LIU, Zhenyi XU, Yu KANG, Yang CAO, Yunbo ZHAO, 2025. Multisensor contrast neural network for remaining useful life prediction of rolling bearings under scarce labeled data. *Front Inform Technol Electron Eng*, 26(7):1180-1193.
https://doi.org/10.1631/FITEE.2400753

1 Analysis of dataset

Fig.S1 shows the time–frequency matrices of the horizontal and vertical acceleration of bearing 1_1 under different health states. The differences not only between healthy and faulty states but also between horizontal and vertical accelerations can be displayed. This is because the wavelet transform mines the hidden frequency domain information in the waveform, which makes the displayed information more abundant.

2 Extended experimental analysis

2.1 Baselines

Due to our goal of RUL prediction under scarce labeled data, we choose six supervised methods to better show the method's performance.

CABLSTM (Luo and Zhang, 2022): A convolution-based attention mechanism bidirectional long and short-term memory network (CABLSTM) is proposed for end-to-end lifetime prediction of rotating machinery. First, the input signal is passed through the convolutional neural network (CNN) to obtain the feature information. Second, the obtained feature information is input into a bidirectional long and short-term memory network (Bi-LSTM) network with an attention mechanism.

2D-LSTM (Li et al., 2022): A two-dimensional long short-term memory (2D-LSTM) based fusion network for RUL prediction. 2D-LSTM is used to extract the depth-temporal features of sensor data one by one, and an an information fusion unit (IFU) is used to fuse multisensor features to predict the RUL of the bearing.

MSGCNN-TR (Guo et al., 2022): A transformer prediction model with multiscale gated convolutional neural network (MSGCNN-TR) is used to predict the RUL of rolling bearings.

HA-ConvLSTM (Zuo et al., 2023):A hybrid attention-based convolutional long short-term memory (HA-ConvLSTM) method for evaluating RUL of bearings is used. First, the original signal is decomposed using multiple wavelets. Second, the ConvLSTM network can adaptively weight the wavelet coefficient channel. Finally, the learned features are used for evaluating RULs through a multilayer perceptron.

BR-GDAU (Yang et al., 2023): A bidirectional recursive gated dual-attention unit (BR-GDAU) is used to predict the RUL in the accelerated degradation phase. Two attention gates are introduced into the classical gated recurrent unit (GRU) to construct a bi-directional structure to fully learn the forward and backward degradation patterns of the time series and to correct the initial hidden state of the forward network by the final hidden state of the backward network.

CATA-TCN (Lin et al., 2024): A channel attention and temporal attention method based on temporal convolutional network (CATA-TCN) is proposed. The channel attention is integrated into the temporal convolutional network (TCN) to focus on sensor signals that are critical for RUL prediction and suppress signals that are unimportant in the long-term range.

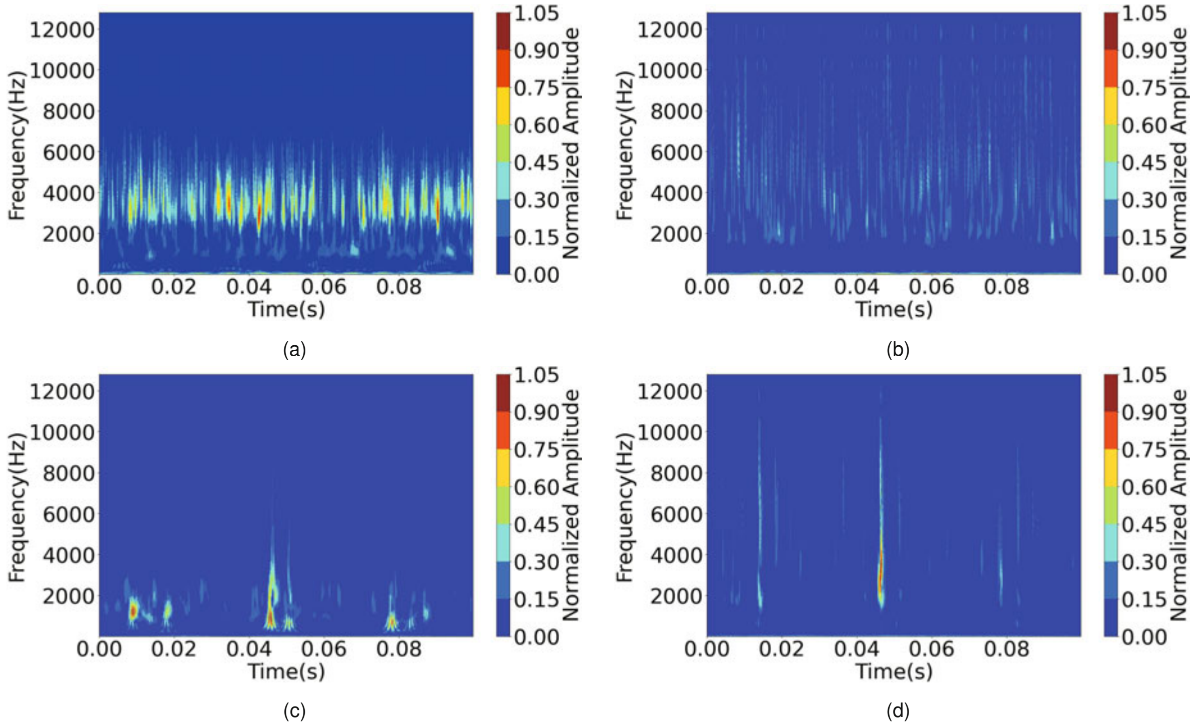


Fig. S1 Results for bearing 1_1: (a) time–Frequency matrix of horizontal acceleration for the first sample; (b) time–Frequency matrix of vertical acceleration for the first sample; (c) time–Frequency matrix of horizontal acceleration for the last sample; (d) time–Frequency matrix of vertical acceleration for the last sample

2.1.1 Experimental fairness

Our problem setting involves massive unlabeled sensor data and scarce sensor data with RUL labels. The supervised method uses training data that are consistent with our method in the finetuning phase. A large amount of unlabeled sensor data is open to every method, it is just difficult for supervised methods to effectively utilize unlabeled sensor data. Moreover, the sensor data with RUL labels used in the finetuning phase of our method are exactly the same as the labeled data used in the supervised method. Thus, all comparisons of our method with the supervised baselines are fair.

2.1.2 Comparison with supervision baselines

Tab. S1 presents the comparative results of the supervised baselines for RUL prediction for bearings with only 50% labeled degradation data and our method. The supervised methods typically assume the availability of a substantial amount of labeled RUL data for accurate prediction. When facing limited labeled RUL data, the best score for the supervised approach is 0.463 and the best MAPE is 0.243, while the worst score for the self-supervised baseline is also 0.525, and the worst MAPE is 0.219. This shows that the supervised baselines tend to be poorer than the self-supervised solutions when the amount of labeled RUL data is limited. Then our proposed approach outperforms the supervised methods significantly, with a minimum decrease of 0.131 in MAPE, and a minimum increase of 0.275 in the score. These results highlight the inherent difficulty of learning reliable degradation patterns from a limited amount of labeled RUL data. Our proposed approach leverages the abundant unlabeled sensor data to obtain the degradation data distribution, facilitating the generation of a well-pretrained model without requiring a large amount of labeled degradation data. This approach contributes to the reduction of the dependence on labeled degradation data.

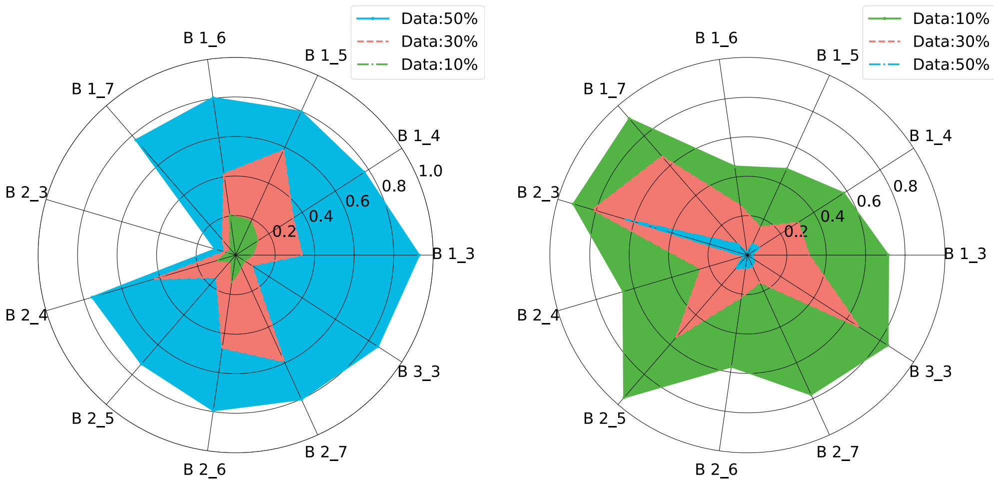
Table. S1 Experimental results compared to supervised methods with 50% labeled data and sufficient unlabeled data

Metrics	Method	B 1_3[1]	B 1_4	B 1_5	B 1_6	B 1_7	B 2_3	B 2_4	B 2_5	B 2_6	B 2_7	B 3_3	Mean
Actual RUL(s)	-	5730	339	1610	1460	7570	7530	1390	3090	1290	580	820	-
Predicted RUL(s)	CABLSTM	5451	460	1214	1043	6626	3045	1468	2757	984	237	752	-
	2D-LSTM	1141	861	2138	3755	1340	2003	376	1614	120	2098	649	-
	MSGCNN-TR	4705	293	988	1135	245	2823	605	2657	669	202	136	-
	HA-ConvLSTM	289	203	285	373	357	3268	1168	2487	1168	608	721	-
	BR-GDAU	5521	127	2115	1511	6586	**3462**	1117	2764	1098	485	542	-
	CATA-TCN	6021	211	1114	1086	6649	2408	1197	2408	1033	511	651	-
	Our method	**5607**	**314**	**1505**	**1483**	**7000**	2670	**1418**	**2807**	**1204**	**543**	**783**	-
A_i	CABLSTM	0.845	0.007	0.426	0.372	0.649	0.127	0.459	0.688	0.440	0.129	0.750	0.445
	2D-LSTM	0.062	0.000	0.011	0.000	0.058	0.079	0.080	0.191	0.043	0.000	0.485	0.092
	MSGCNN-TR	0.538	0.625	0.262	0.462	0.035	0.115	0.141	0.615	0.189	0.104	0.056	0.286
	HA-ConvLSTM	0.037	0.249	0.058	0.076	0.037	0.141	0.575	0.508	0.721	0.512	0.658	0.325
	BR-GDAU	0.881	0.114	0.013	0.616	0.637	**0.154**	0.506	0.694	0.597	0.567	0.309	0.463
	CATA-TCN	0.495	0.270	0.344	0.412	0.656	0.095	0.618	0.465	0.501	0.662	0.490	0.455
	Our method	**0.928**	**0.774**	**0.798**	**0.804**	**0.770**	0.107	**0.756**	**0.728**	**0.794**	**0.802**	**0.855**	**0.738**
MAPE	CABLSTM	0.049	0.357	0.246	0.286	0.125	0.596	0.056	0.108	0.237	0.591	0.083	0.248
	2D-LSTM	0.801	1.540	0.328	1.572	0.823	0.734	0.729	0.478	0.907	2.617	0.209	0.976
	MSGCNN-TR	0.179	0.136	0.386	0.223	0.968	0.625	0.565	0.140	0.481	0.652	0.834	0.471
	HA-ConvLSTM	0.950	0.401	0.823	0.745	0.953	0.566	0.160	0.195	0.095	0.048	0.121	0.460
	BR-GDAU	0.036	0.625	0.314	0.035	0.130	**0.540**	0.196	0.106	0.149	0.164	0.339	0.239
	CATA-TCN	0.051	0.378	0.308	0.256	0.122	0.680	0.139	0.221	0.199	0.119	0.206	0.243
	Our method	**0.021**	**0.074**	**0.065**	**0.016**	**0.075**	0.645	**0.020**	**0.092**	**0.067**	**0.064**	**0.045**	**0.108**

[1] B denotes bearing, e.g. B 1_3 denote bearing 1_3.

2.2 Study on the effectiveness of scarce labeled data

To investigate the effect of the amount of labeled RUL data in the finetuning phase on the model performance, we conduct experiments with different amounts of labeled RUL data, including 10%, 30%, and 50%. We keep the data used in the alternate contrast phase the same for all experiments. Fig. S2 demonstrates the model's performance on 11 bearings under different amounts of fine-tuning of data. As the quantity of labeled data increases in the fine-tuning phase, the score gradually increases while the MAPE gradually decreases. It can be observed that when only 10% of the labeled data is used, the model's performance deteriorates significantly, making it difficult to effectively predict the remaining useful life [RUL].



(a) Score of our method under different percentages of labeled data.

(b) MAPE of our method under different percentages of labeled data.

Fig. S2 Metrics of our method under different percentages of labeled data

References

Guo, D., Cao, Z., Fu, H., Li, Z., 2022. Remaining useful life estimation for rolling bearings using msgcnn-tr. IEEE Sensors Journal 22 (24), 24333–24343.

Li, Y., Wang, H., Li, J., Tan, J., 2022. A 2-d long short-term memory fusion networks for bearing remaining useful life prediction. IEEE Sensors Journal 22 (22), 21806–21815.

Lin, L., Wu, J., Fu, S., Zhang, S., Tong, C., Zu, L., 2024. Channel attention & temporal attention based temporal convolutional network: A dual attention framework for remaining useful life prediction of the aircraft engines. Advanced Engineering Informatics 60, 102372.

Luo, J., Zhang, X., 2022. Convolutional neural network based on attention mechanism and bi-lstm for bearing remaining life prediction. Applied Intelligence 52 (1), 1076–1091.

Yang, L., Liao, Y., Duan, R., Kang, T., Xue, J., 2023. A bidirectional recursive gated dual attention unit based rul prediction approach. Engineering Applications of Artificial Intelligence 120, 105885.

Zuo, T., Zhang, K., Zheng, Q., Li, X., Li, Z., Ding, G., Zhao, M., 2023. A hybrid attention-based multi-wavelet coefficient fusion method in rul prognosis of rolling bearings. Reliability Engineering & System Safety 237, 109337.