PCC2023
XXI Polish Control Conference

POLITECNICO MILANO 1863

DIPARTIMENTO DI MECCANICA

Intelligent Fault Diagnosis for Rotating Machinery

Hamid Reza Karimi
Professor, Politecnico di Milano
hamidreza.karimi@polimi.it

mecc

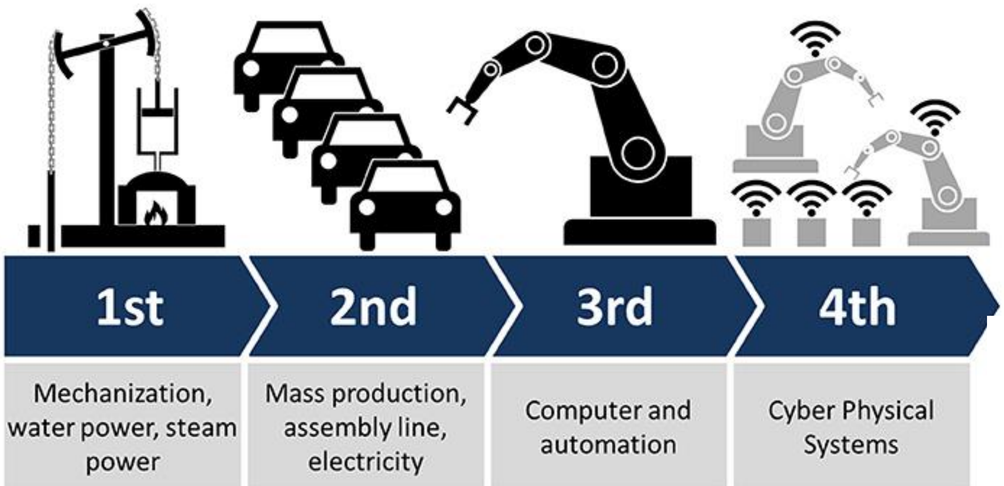
Dipartimento Di Meccanica
Via La Masa, 1 - 20156 Milano (Campus Bovisa)
peccmecc@cert.polimi.it

Gliwice, 26/06/2023

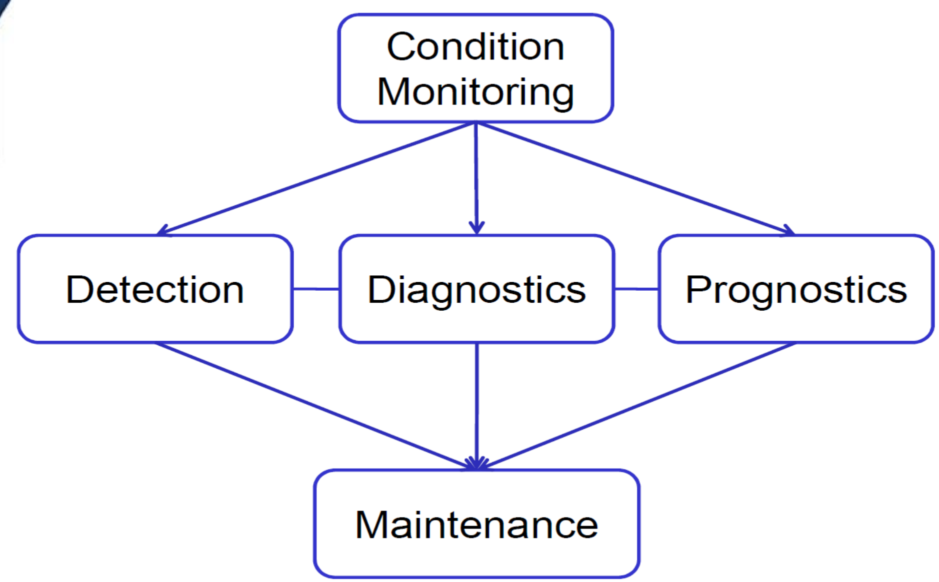
Outline

Outline

Revolutions in Industry



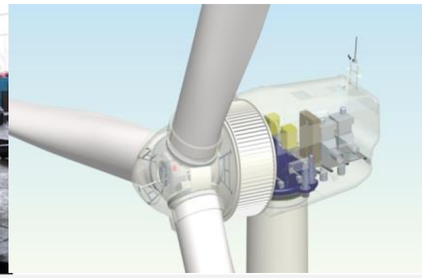
@ Roser, C. (2015). Industry 4.0 Wikipedia



Faults in rotary machinery

For major/key equipment in the fields of iron and steel, energy, transportation, aerospace, etc., the safety and security of its service operation is very important.


Strip Mill


Wind Turbines


Express train


Aircraft engine

CNC machine


Industrial robot


Helicopter


Automobile

Carrying out equipment fault diagnosis and predictive maintenance has important research significance and engineering application value to ensure production safety.

Faults in rotary machinery

Key equipment –> rotating parts –> key components: gears, bearings, and rotor systems are the main failure components.

Gear drive

Bearing

Rotor

blade

Bearing cage damage

Motor failure

Turnout machine failure

Engine failure

The monitoring and fault diagnosis of core components has always been a research hotspot in the field of industrial fault diagnosis.

Health Management System

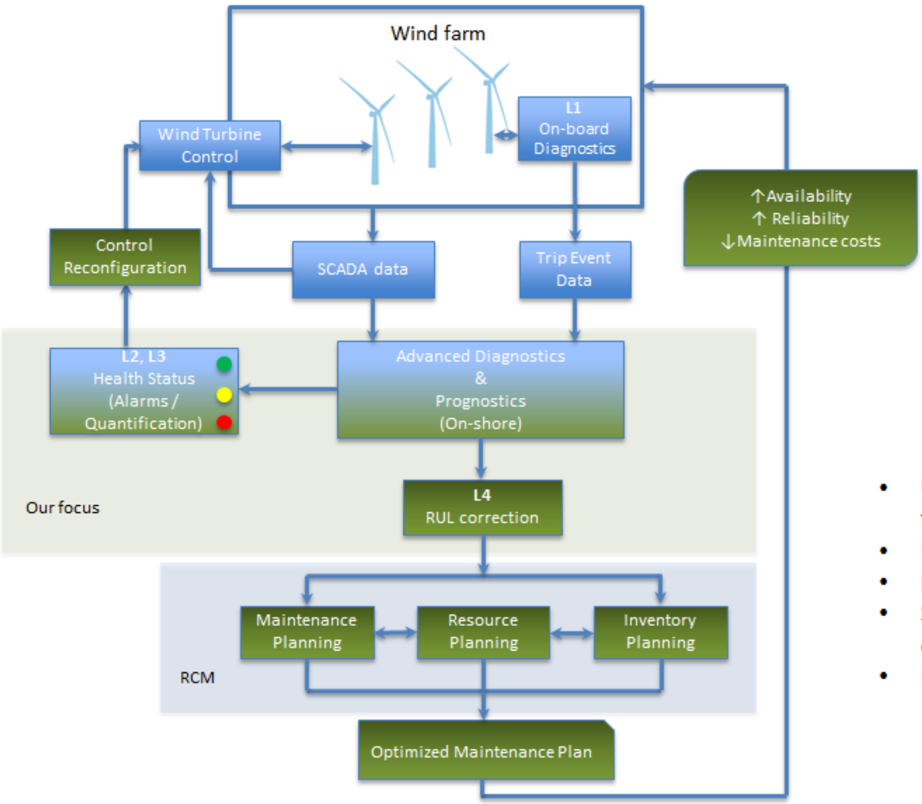
Health Management helps in:

- **Incipient failure detection** – detecting failures even before they substantially effect the performance
- **Prevention of fault progression** – eradicating fault conditions before secondary faults develop
- **Prediction of progression from fault to failure** – accurate prognosis for remaining useful life
- **Efficient maintenance planning** – economize on maintenance efforts, ensure best availability
- **Feedback to control laws** – Modify control laws based on the current health condition to extend life while obtaining the best possible performance.

Overall System Health for increased RAM (Reliability and Maintainability) and minimized O&M (Operation and Maintenance) costs

Health Management for Wind farms

Towards Farm-level Health Management

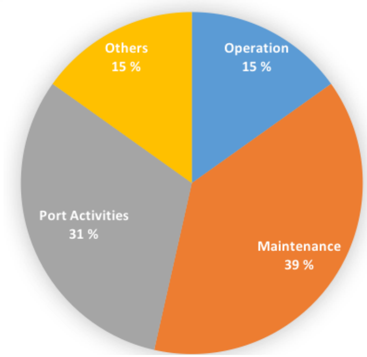


Health Monitoring

- Level 1 Detection
- Level 2 Isolation
- Level 3 Quantification
- Level 4 Prognosis

Present Future

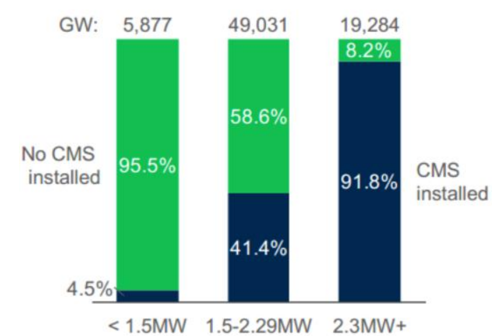
O&M Costs: Offshore Wind Farms



Benefits

- Unprecedented insight into fleet health
- Less maintenance costs
- No "on-failure" repairs
- Short & responsive supply chain
- Minimum inspections

Penetration of CMS – US market



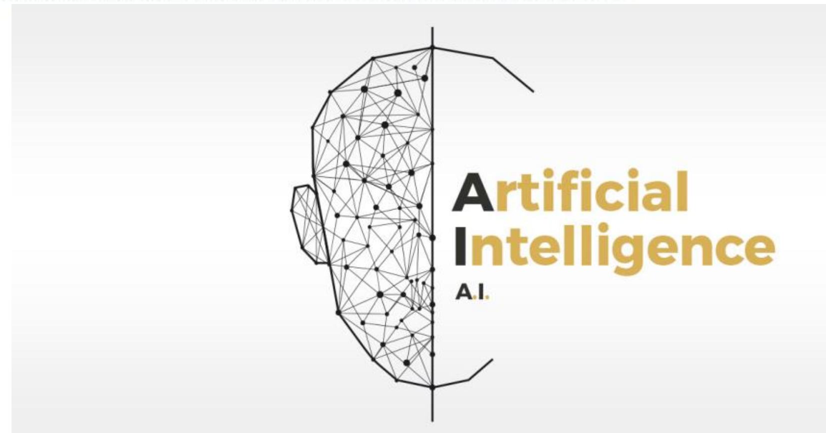
Source: MAKE
Note: EOY 2015

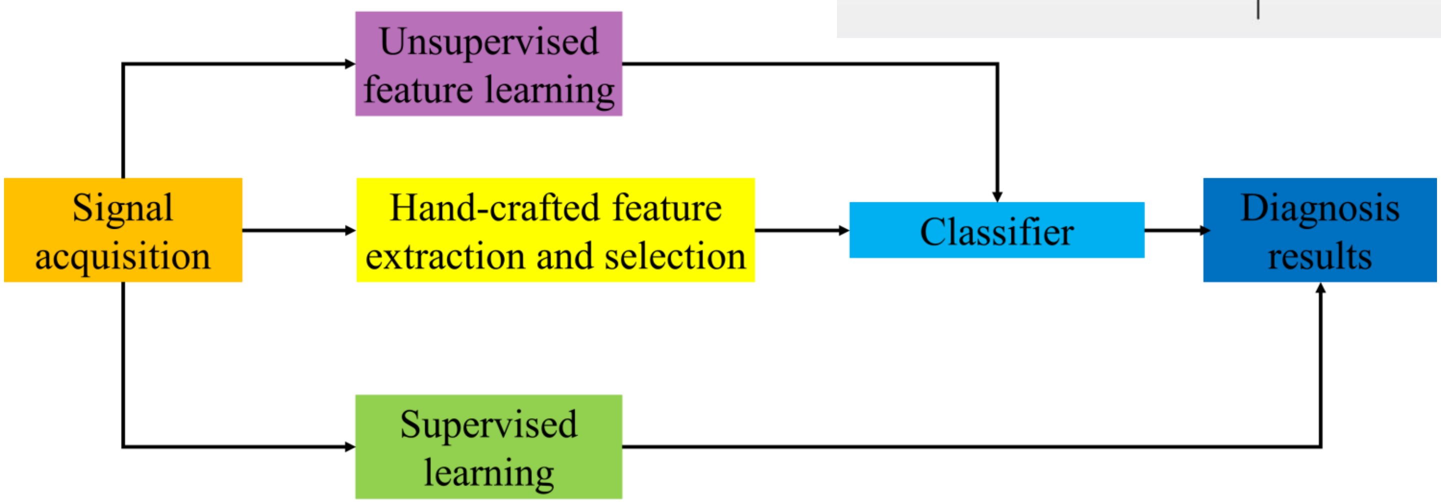
Health Management – capability to make intelligent, informed, appropriate decisions about maintenance and logistics actions based on diagnostics/prognostics information, available resources and operational demand. – Definition, JSF Program

Intelligent fault diagnosis Framework

General intelligent fault diagnosis framework

- ❑ Unsupervised learning
- ❑ Supervised learning
- ❑ Reinforcement learning

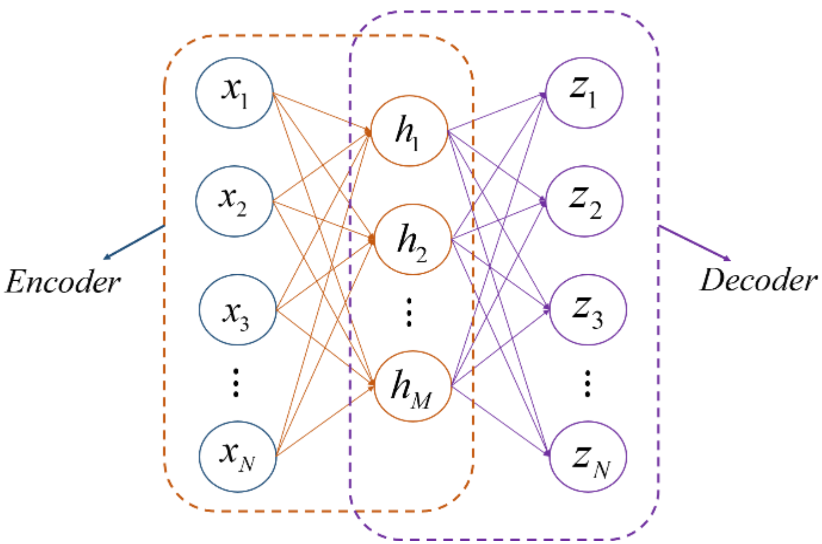




Three general intelligent fault diagnosis frameworks (Sensors, 2017)

Unsupervised learning: Auto-encoder

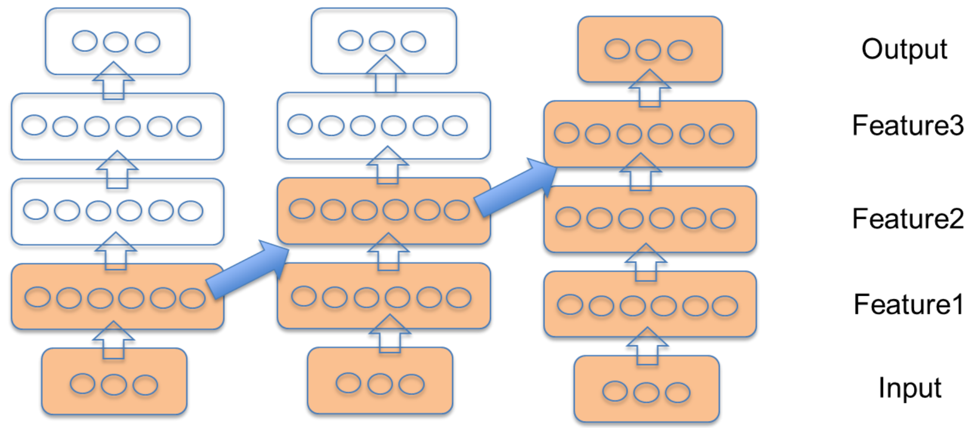
$$h = f\left(W^{(1)}x + b^{(1)}\right)$$



The structure of standard Auto-encoder (Knowledge-Based Systems, 2020)

Advantages: Computation speed

Disadvantages: High-dimensional signal



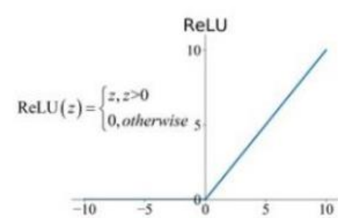
The training process of Stacked Auto-encoder (MSSP, 2018)

Advantages: Training efficiency

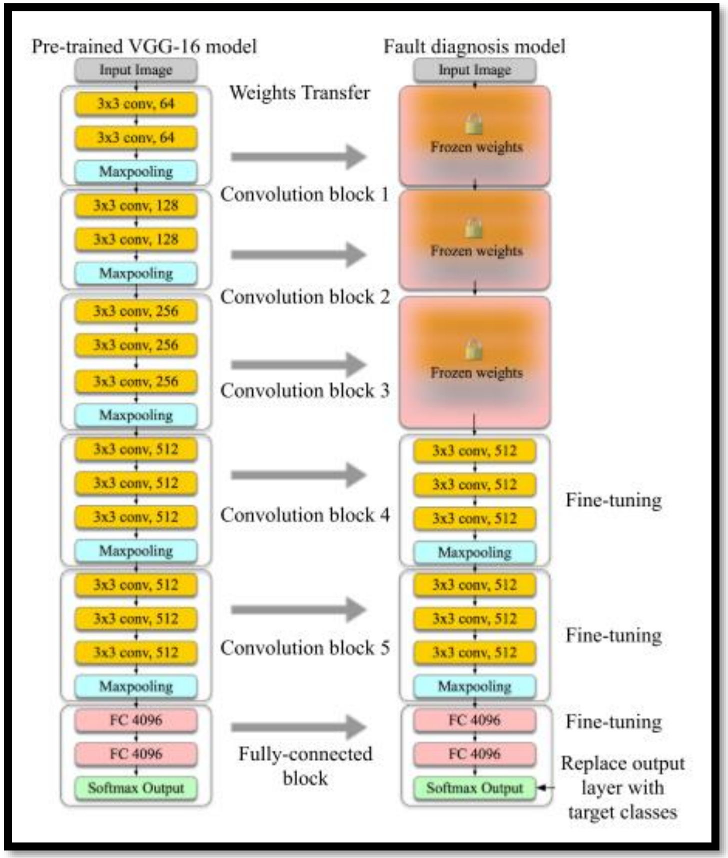
Disadvantages: High-dimensional signal

$$f(x) = \max(0, x)$$

Rectified Linear Unit (ReLU) function



Supervised learning: Convolutional Neural Networks



VGG-16 based on Transfer Learning

Advantages:

- ➢ High computation speed
- ➢ High accuracy

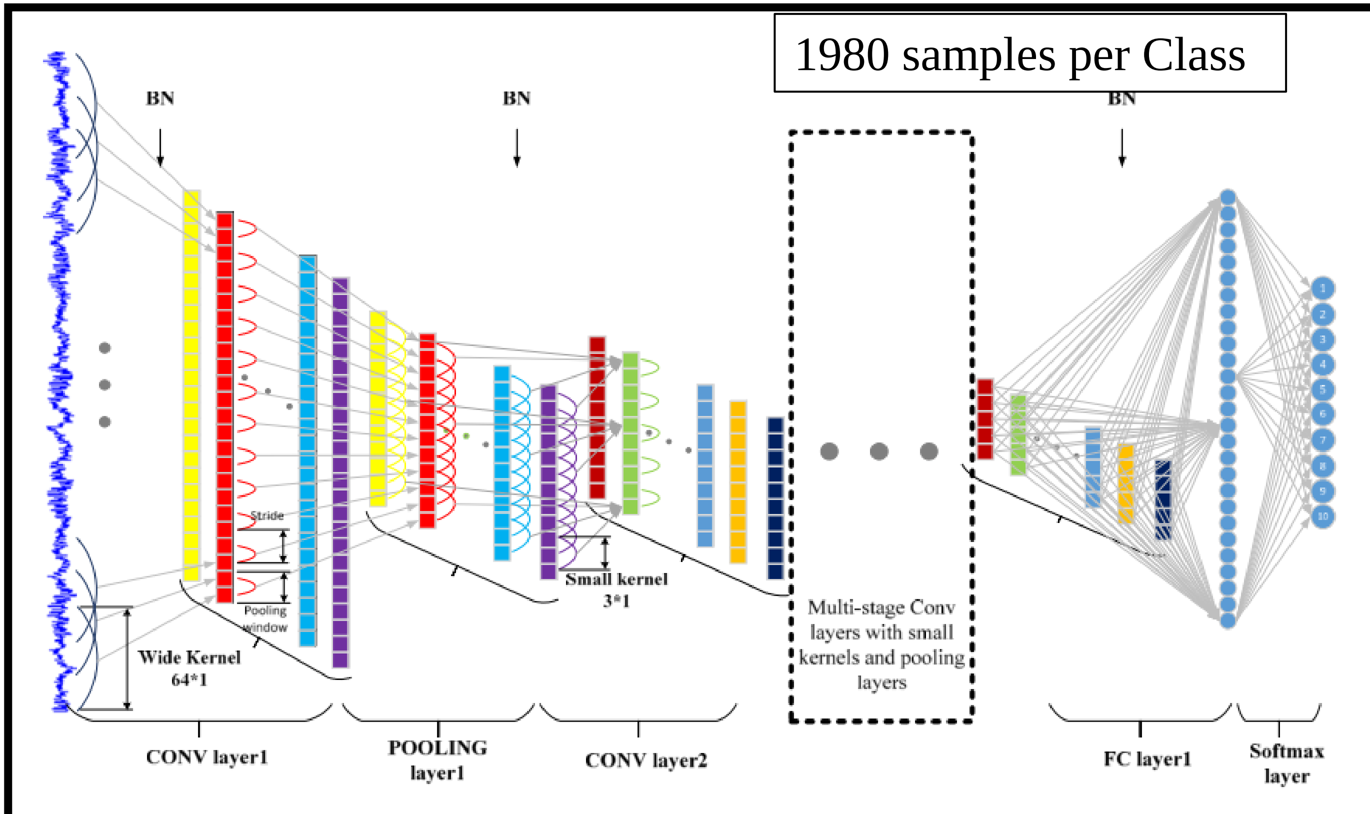
Disadvantages:

- ➢ Need to transform the raw vibration signals
- ➢ Need more than 500 samples per Class

500 samples per Class

Transfer learning procedure trains the three highest-level blocks of the pretrained VGG-16 network (TIE, 2018)

Supervised learning: Convolutional Neural Networks



Architecture of the WKDCNN model (Sensors, 2017).

Wide-kernel deep convolutional neural networks

Advantages:
- Without signal processing

Disadvantages:
- General accuracy
- Need 1980 samples per Class

Outline

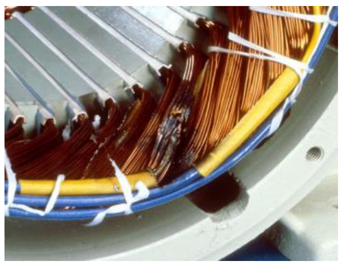
02 Intelligent Filter-based Fault Analysis

05 Multi-source information fusion

Electrical and Mechanical faults

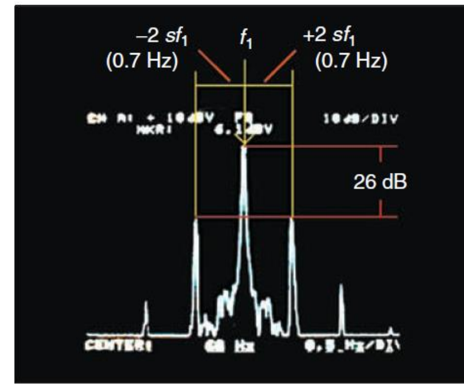
Manifest as periodic disturbances in supply current

Electrical faults

Mechanical faults

Variations in electric circuit/ airgap magnetic field



- Bearing fault $f_v \approx 0.6nf_r, 0.4nf_r,\ f_{brg} = f_s \pm mf_v$
- Stator winding fault $f_{st} = f_s[\frac{n}{P}(1-s) \pm k]$
- Air gap eccentricity $f_{age} = f_s[(R \pm n_d)\left(\frac{1-s}{P}\right) \pm n_{ws}]$
- Broken rotor bar $f_{brb} = f_s(1 \pm 2s)$

➢ f_r: the rotor rotational frequency;
➢ f_s : the supply frequency
➢ f_v : the outer raceway fault frequency

Fault models

Detailed induction motor modeling based on modified winding function theory (MWFTh)

$$\{V_s\}_{(3\times1)} = [R_s]_{(3\times3)}\{i_s\}_{(3\times1)} + \frac{d}{dt}\left([L_s]_{(3\times3)}\{i_s\}_{(3\times1)} + [L_{sr}]_{3\times N_r}\{i_r\}_{N_r\times1}\right)$$

$$\{0\}_{(N_r\times1)} = [R_r]_{(N_r\times N_r)}\{i_r\}_{(N_r\times1)} + \frac{d}{dt}\left([L_r]_{(N_r\times N_r)}\{i_r\}_{(N_r\times1)} + [L_{rs}]_{N_r\times3}\{i_s\}_{3\times1}\right)$$

The inductances are calculated considering the modified airgap function, due to faults.

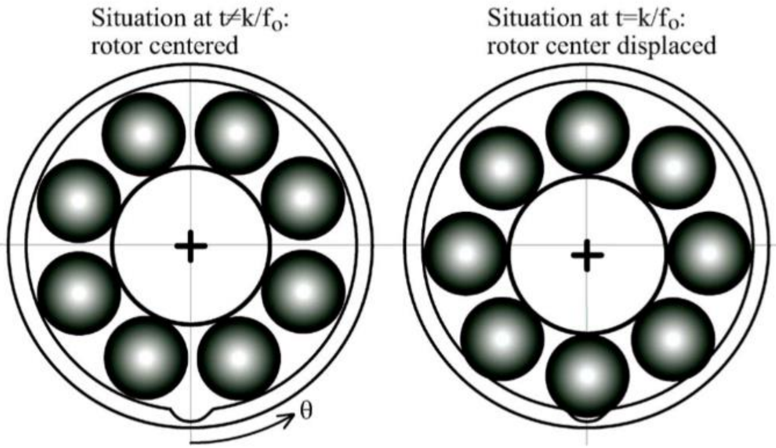
$$L_{AB} = \mu_0 r l \int_0^{2\pi} M_A(\phi,\theta) n_B(\phi,\theta) g^{-1}(\phi,\theta)\, d\phi$$

Air gap eccentricity

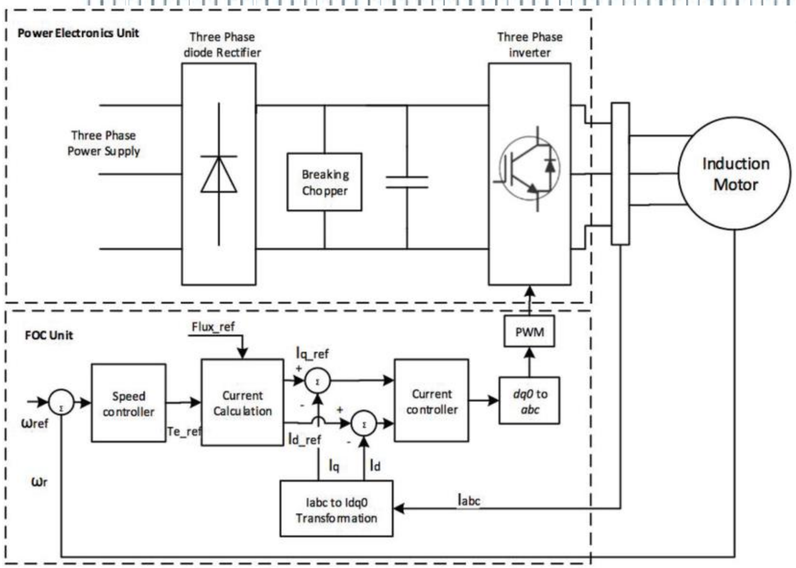
$$g(\phi,\theta) = g_0(1 - \delta\cos(\phi - \theta))$$

Bearing fault (BRG)

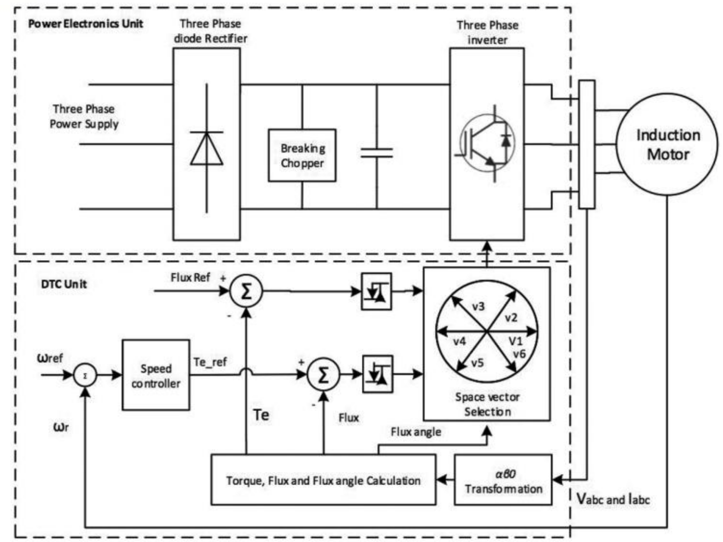
$$g(\phi,\theta,t) = g_0[1 - e_0\cos(\phi)\sum_{k=-\infty}^{+\infty}\delta\left(t - \frac{k}{f_c}\right)]$$



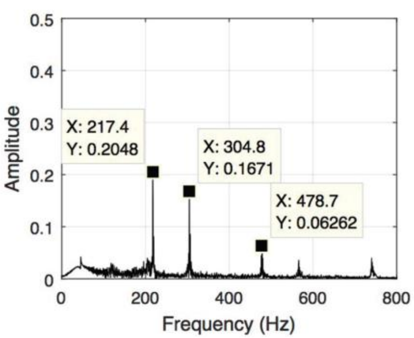
Closed-loop motor fault diagnostics



Field-oriented control system.

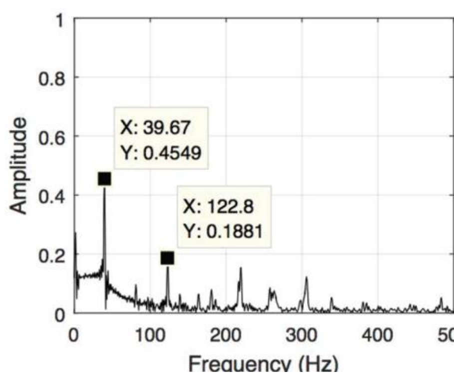


Direct torque control system.

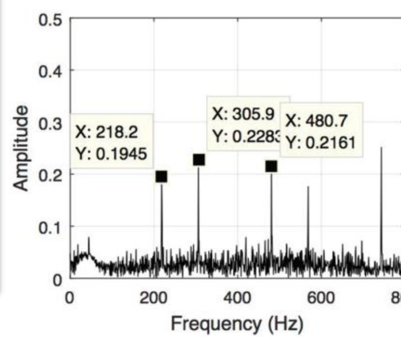


Current spectrum of a healthy motor with FOC

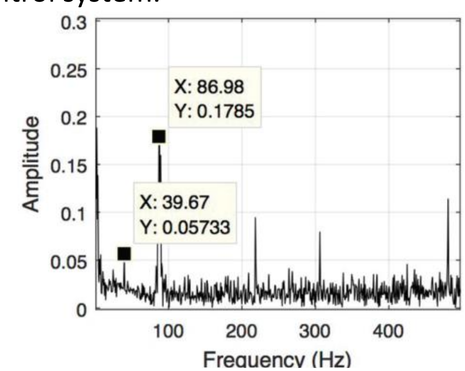
$f_s = 43.49$ Hz The supply frequency.



Fault frequencies at f_{AGE_1} with FOC.



Current spectrum of a healthy motor with DTC.



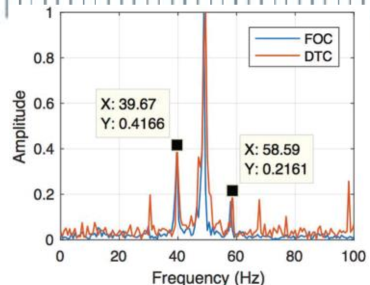
Fault frequencies at f_{AGE_1} with DTC.

Closed-loop motor fault diagnostics

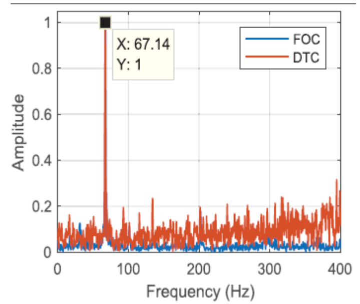
(a) broken rotor bars

(b) stator-turn fault

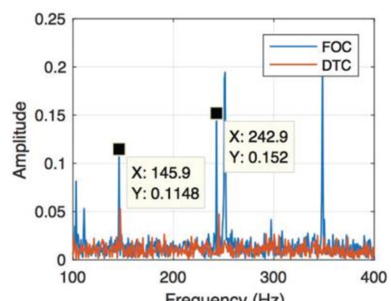
(c) bearing outer race fault.



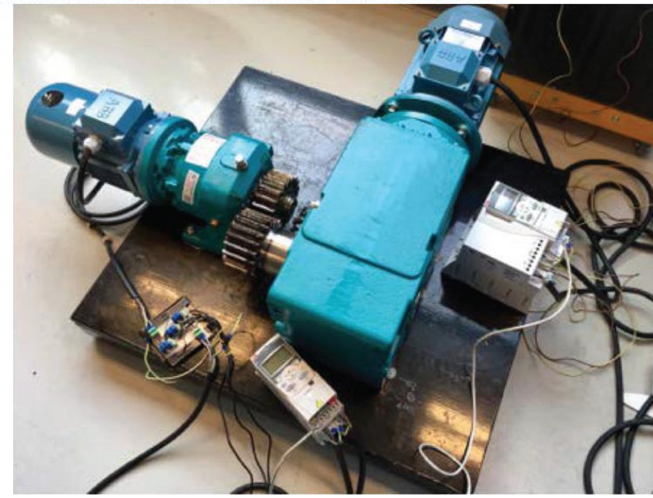
Single-line current spectrum with BRB fault at 1420 r/min



i_P current spectrum with STF at 1420 r/min.



Single-line current spectrum with BRG fault at 1420 r/min.



Fault frequencies in closed-loop operation of the test motor in the laboratory set-up.

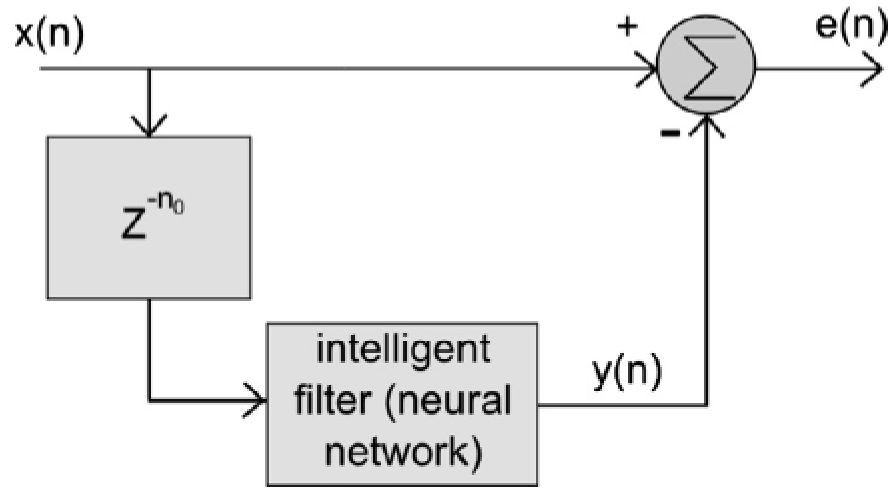
Control	Speed (r/min)	f_{BRB} (Hz)	f_{BRG} (Hz)
FOC	1000	27.5, 29.8	85.6, 94.4
		32.8, 41.5	145.6, 154.4
FOC	1420	36, 41	121.8, 133.8
		56.5, 61.2	207, 219
DTC	1000	28.03, 31.5	84.9, 95.1
		31.6, 42.1	144.9, 155.1
DTC	1420	25.1, 31.4	121.4, 134.2
		39.5, 59.6	206.6, 219.4

FOC: field-oriented control; DTC: direct torque control.

Bearing Fault Diagnosis & Classification

❖ Design of a removing non-bearing fault component (RNFC) filter based on neural networks.

Intelligent RNFC filter design

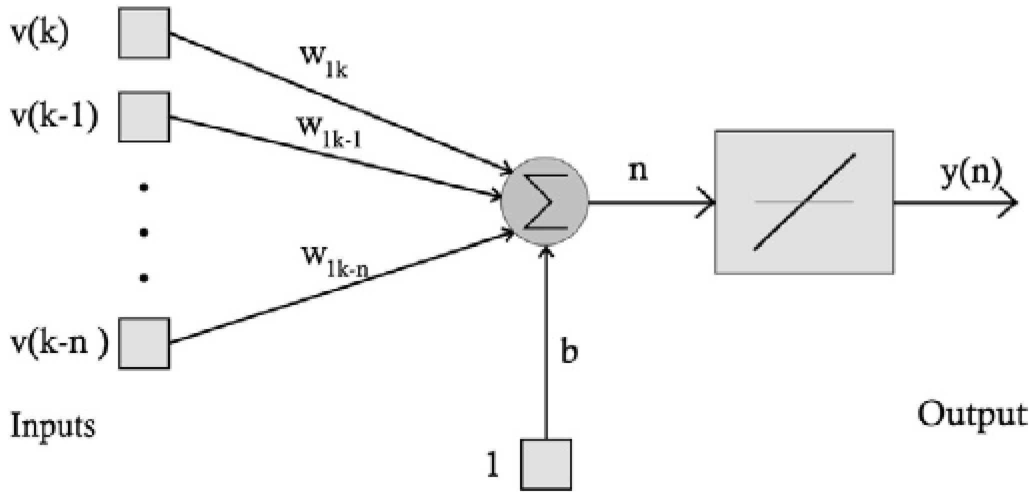


$x(n)$: Motor vibration signal.
$y(n)$: Estimated irrelevant part of the vibration signal (non-bearing fault components).

$e(n)$: Faulty part of the vibration signal.
n_0: Number of data samples.

ADALINE Network

Adaptive Linear Neuron (ADALINE) neural network with purelin activation functions



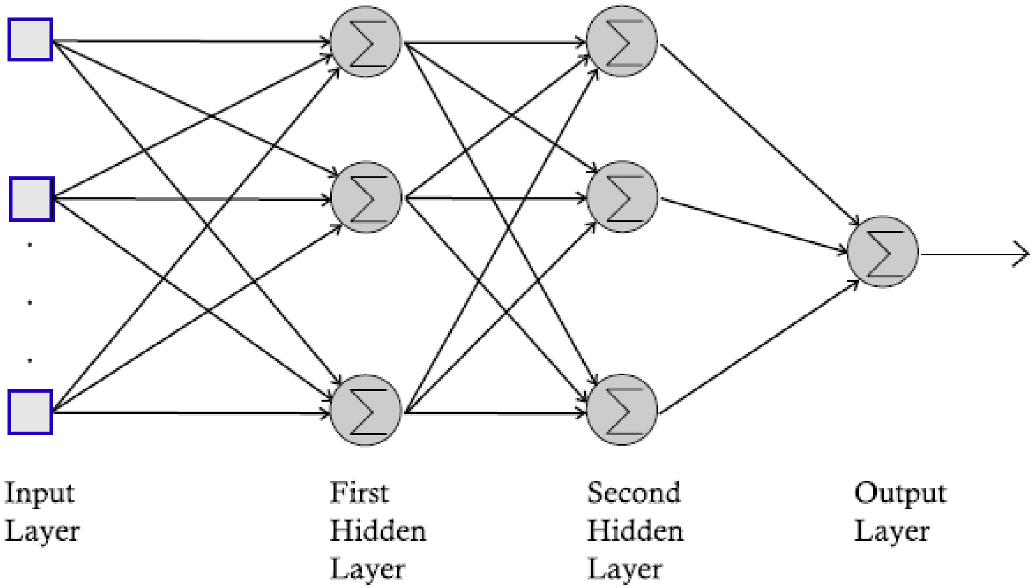
$$p = \begin{bmatrix} healthy(k) \\ healthy(k-1) \\ \vdots \\ healthy(k-n_0) \end{bmatrix} \quad n_0 = 1, 2, \ldots$$

$$t = [healthy(k)] \quad k = 1, 2, \ldots$$

healthy(k) is the sampled vibration signal of a healthy induction motor (k is the indices for the number of samples).

Bearing Fault Diagnosis & Classification

Fault classification based on pattern recognition for healthy and defective bearings in four categories, including healthy condition, inner race defect, outer race defect and double holes in outer race.

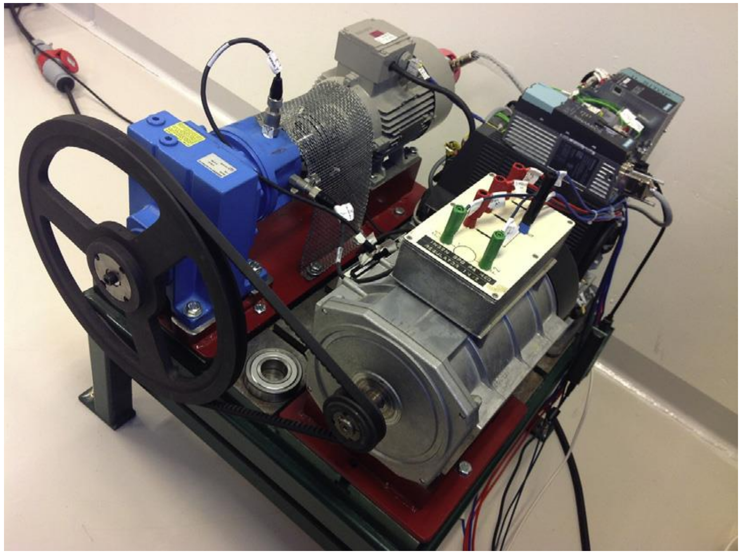


The structure of a multi-layer perceptron network.

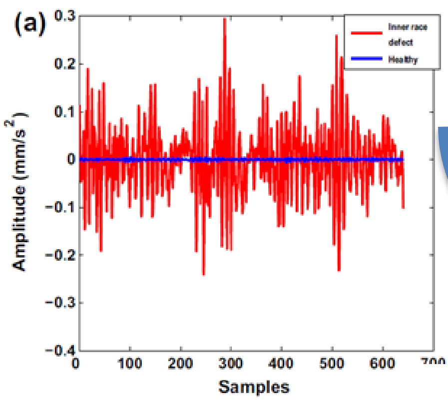
Experimental setup and results

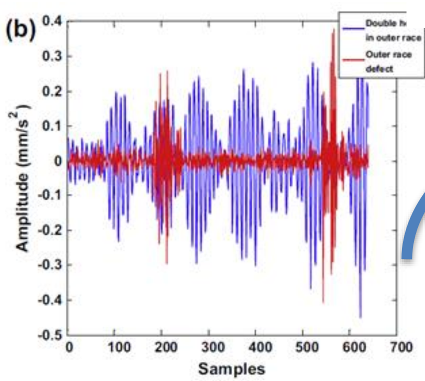
❑ Characterization

- A three-phase, 1.2 kW, 380 V, 1500 rpm, four pole induction motor is used to collect experimental data.
- Both shaft-end and fan-end bearings are 6205-2Z.
- The vibration signal is sampled by Advantech PCI-1711 data acquisition card with 32 kHz sampling frequency using B&K 4395 accelerometer.

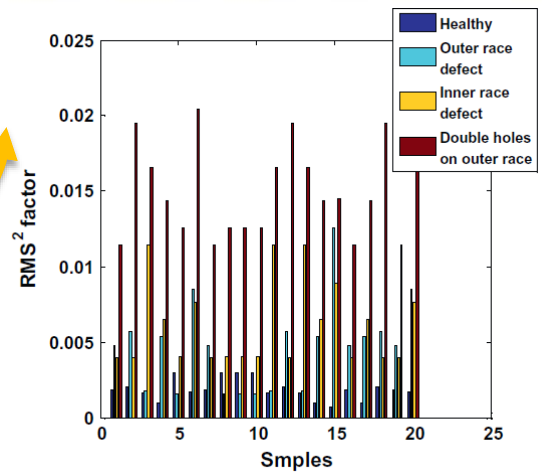


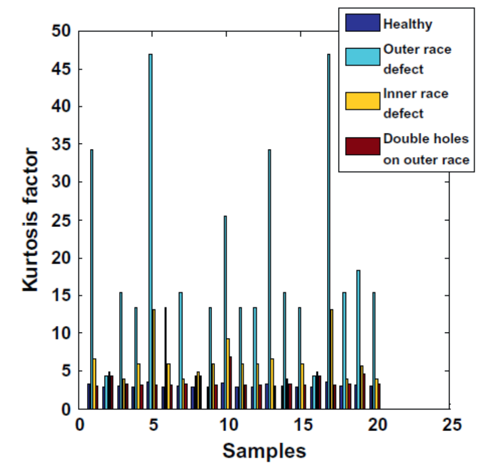
Time-domain features of test data

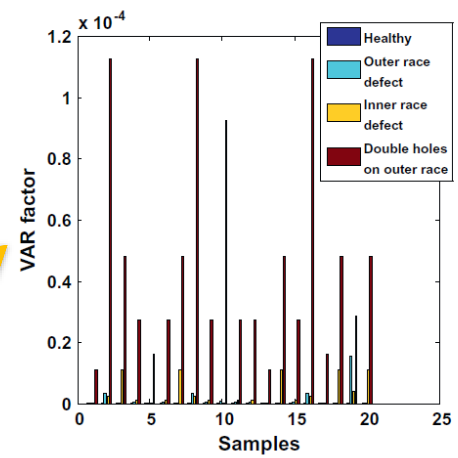


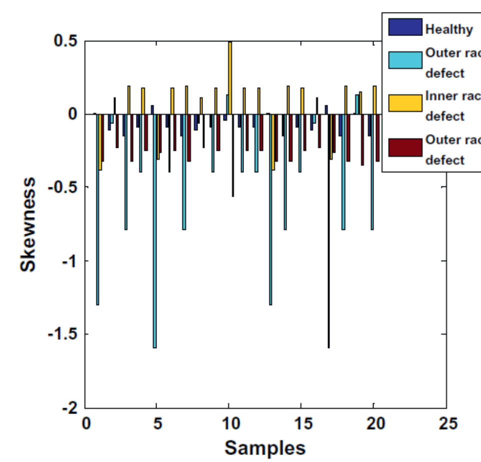


$$p = \begin{bmatrix} \textit{first central moment } (RMS^2) \\ \textit{second central moment } (\textit{variance}) \\ \textit{third central moment } (\textit{skewness}) \\ \textit{forth central moment } (\textit{kurtosis}) \end{bmatrix}$$









RNFC performance

Table 1
Fault detection using RNFC filter.

Net number	Neurons number	Correct classification percent			
		Healthy (%)	Inner race defect (%)	Outer race defect (%)	Double holes in outer race (%)
1	[4 3 2]	100	100	100	100
2	[4 8 2]	100	100	100	100
3	[4 3 5 2]	100	100	100	100
4	[4 10 3 5 2]	0	0	0	0

Table 2
Direct fault detection (fault classification without RNFC filter).

Net number	Neurons number	Correct classification percent			
		Healthy (%)	Inner race defect (%)	Outer race defect (%)	Double holes in outer race (%)
1	[4 3 2]	0	16	60	56
2	[4 8 2]	16	28	16	12
3	[4 3 5 2]	32	12	30	64
4	[4 10 3 5 2]	0	0	0	0

Table 4
Fault detection in presence of low-quality sampled signals using RNFC filter.

Net number	Neurons number	Correct classification percent							
		Healthy (%)		Inner race defect (%)		Outer race defect (%)		Double holes in outer race (%)	
		With filter	Without filter	With filter	Without filter	With filter	Without filter	With filter	Without filter
1	[4 3 2]	100	4	96	0	80	34	100	44
2	[4 8 2]	100	24	100	20	100	78	100	0
3	[4 3 5 2]	100	72	100	8	96	48	100	24

Outline

03 Residual wide-kernel deep convolutional autoencoder

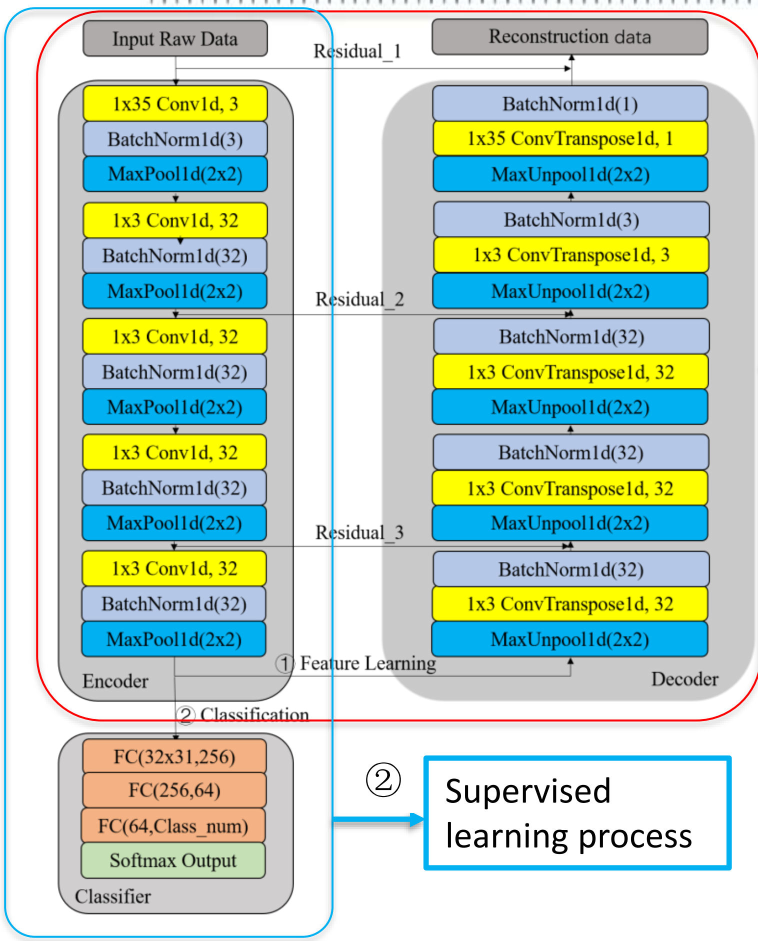
05 Multi-source information fusion

SAE vs CNN

There are several problems with the Standard Auto-encoder and Convolutional Neural Networks.

- To get high accuracy, the input vibration signals are needed to transform into other kinds of signals.
- The feature extraction ability of Standard Auto-encoder to deal with high-dimensional data is not good.
- The traditional convolutional neural networks are easy to over-fitting and gradient vanishing based on the limited data.

Residual wide-kernel deep convolutional auto-encoder (RWKDCAE)



$$\tilde{x} = f\left(W^{(2)}h + b^{(2)}\right)$$

① Unsupervised learning process

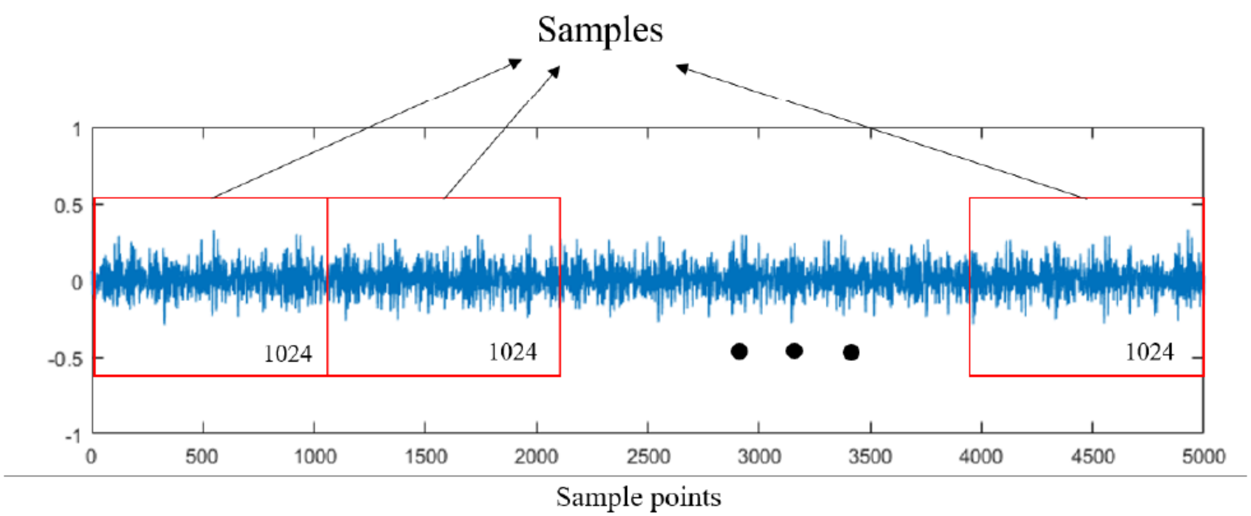
The objective of AE is to minimize

$$Loss = |x - \tilde{x}|^2$$

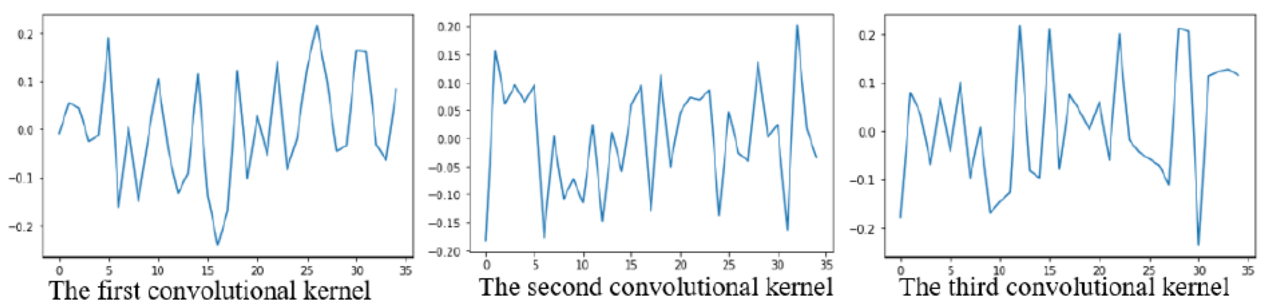
- The wide-kernel convolutional layer is introduced in the convolutional auto-encoder that can ensure the model can learn effective features from the data without any signal processing.
- The residual learning block is introduced in convolutional auto-encoder that can ensure the model with sufficient depth without gradient vanishing and overfitting problems.
- Convolutional auto-encoder can learn constructive features without massive data.

Bearing Fault Diagnosis & Classification

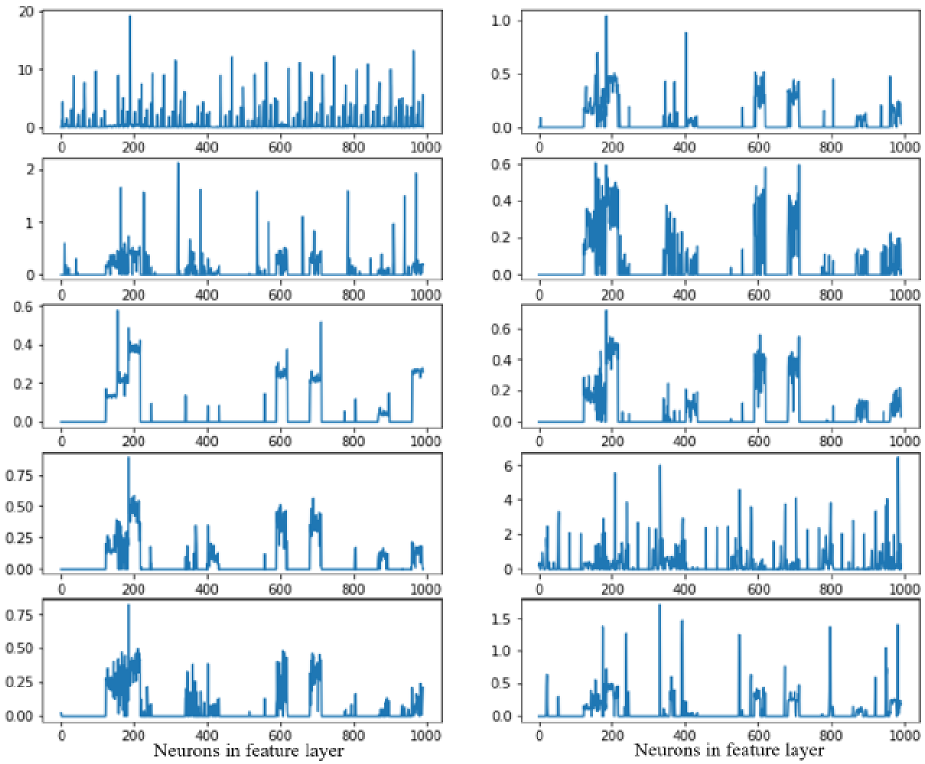
❖ Design of residual wide-kernel deep convolutional autoencoder (RWKDCAE)



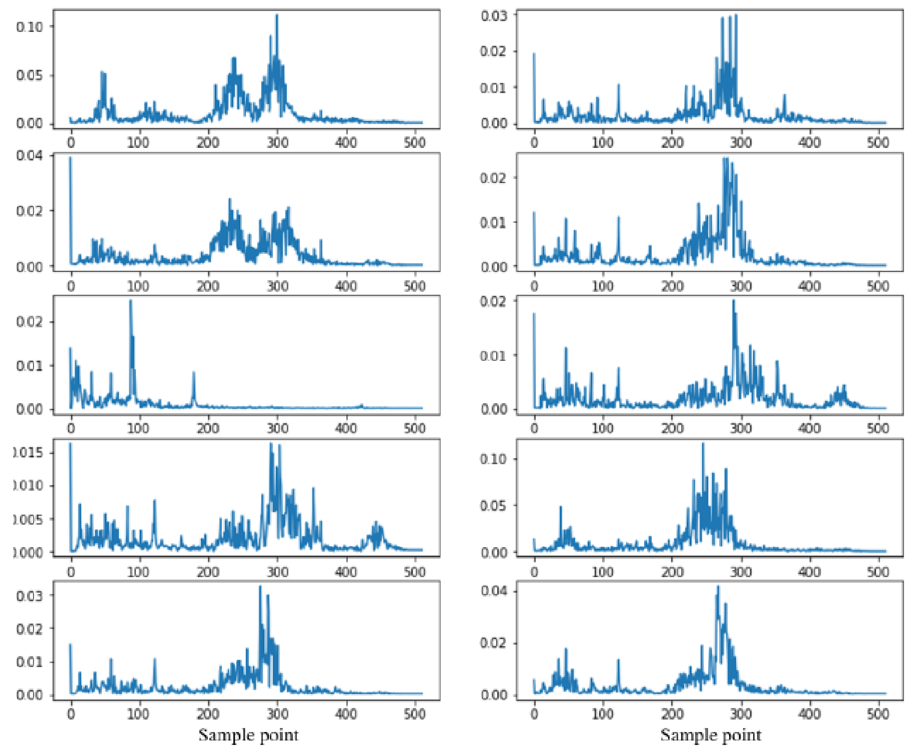
❖ Kernel visualization in the first convolutional layer.



Feature Visualization



Visualization of features from encoder of RWKDCAE with unsupervised learning

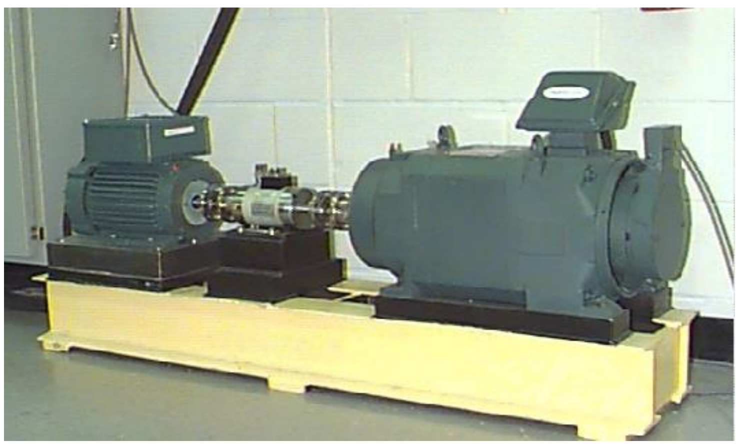


Visualization of features from Frequent-domain signal transform by FFT

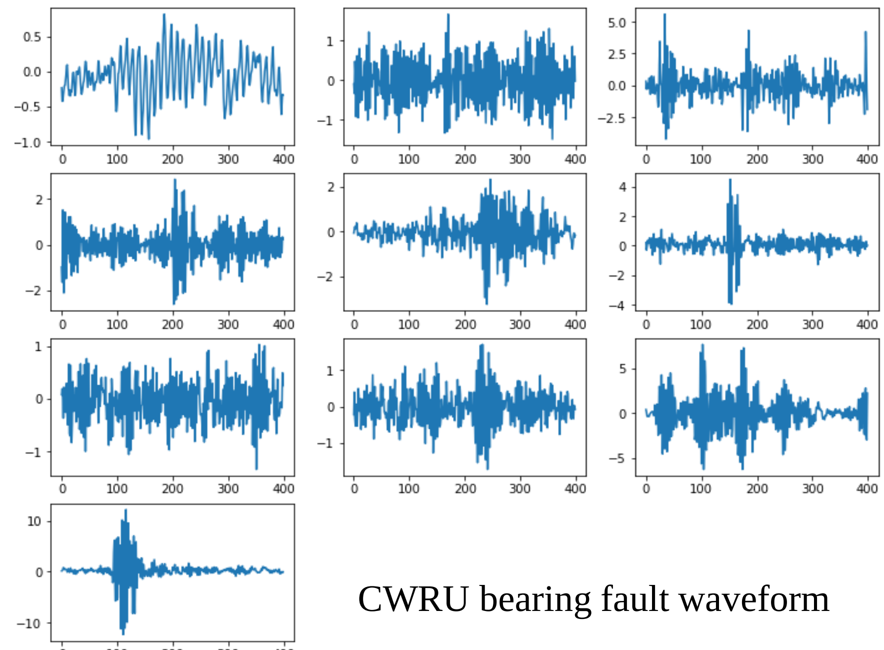
Bearing Dataset

Case Western Reserve University (CWRU) Bearing Fault Dataset

- CWRU bearing dataset is made up of nine fault categories and one normal condition.
- The nine types of faults are divided into three main types, which is the inner raceway fault, the outer raceway fault and the ball fault.
- There are three fault diameters for each fault type, which are 0.007 inches, 0.014 inches and 0.021 inches.



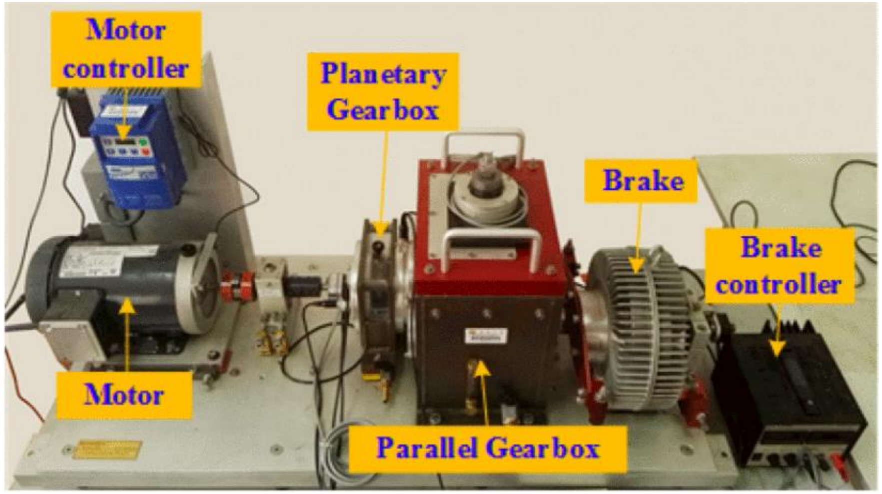
CWRU bearing test rig

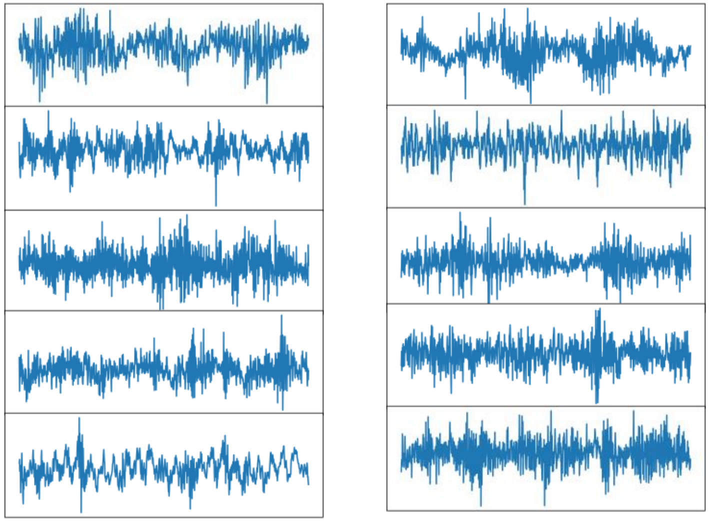


CWRU bearing fault waveform

Gearbox Dataset

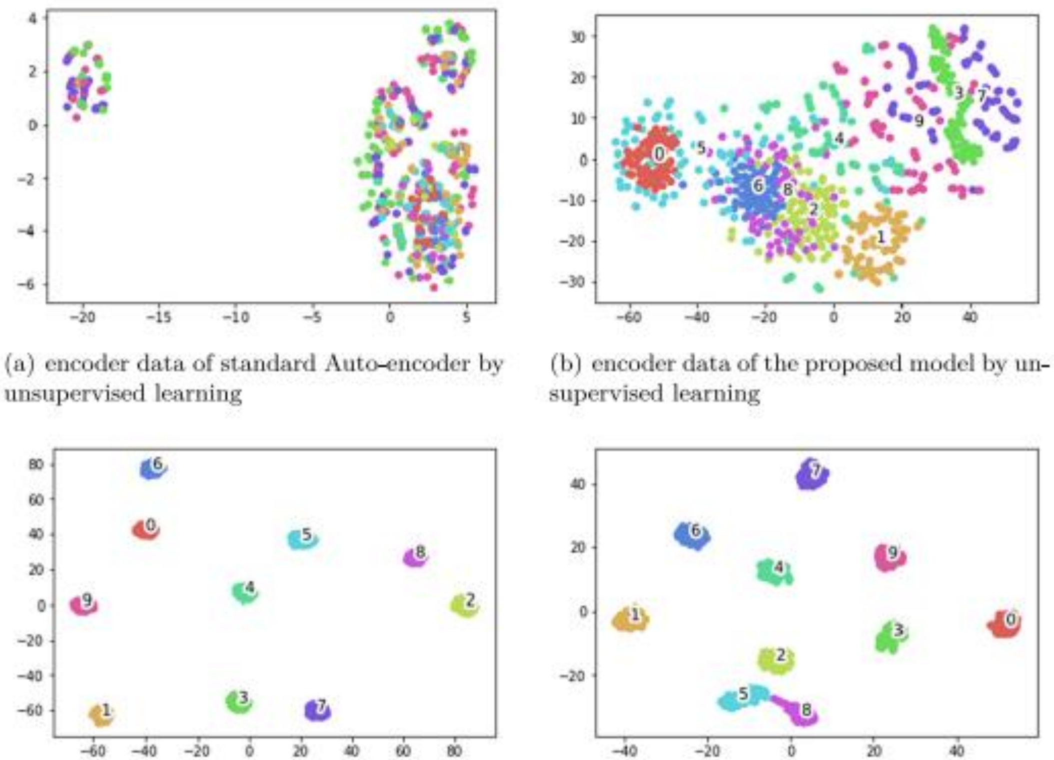
- This gearbox dataset was provided by Southeast University (SEU).
- There were two working conditions in this dataset for the bearing data and gearbox data: 20 HZ–0 V and 30 HZ–2 V.
- The gearbox fault diagnosis for bearing or gear dataset is a 5-class problem (four failure types and one health state), and combining all data together became a 10-class problem.





SEU bearing and gear fault waveform

Feature learning ability comparison



(a) encoder data of standard Auto-encoder by unsupervised learning

(b) encoder data of the proposed model by unsupervised learning

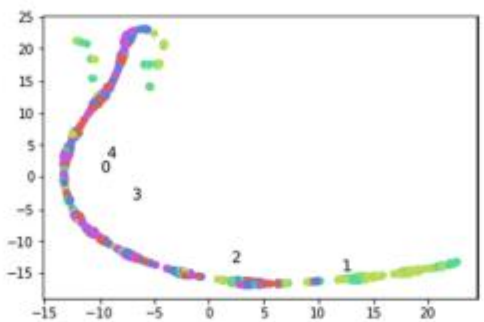
(c) encoder data of the proposed model by supervised learning

(d) frequency-domain data transform from FFT

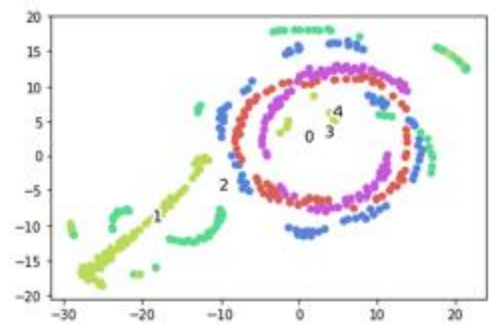
Feature visualization of CWRU bearing dataset

a) the feature that learned by a Standard auto-encoder by unsupervised learning process.
b) the feature that learned by the proposed model by unsupervised learning process.
c) the feature that learned by the proposed RWKDCAE model by supervised learning process.
d) the frequency-domain signals.

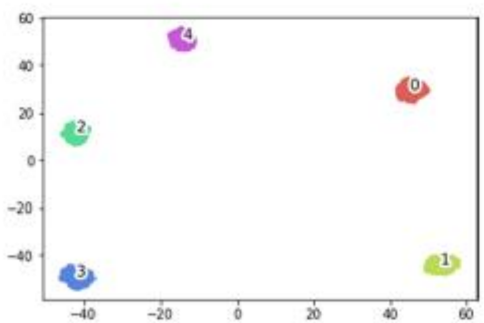
Feature learning ability comparison



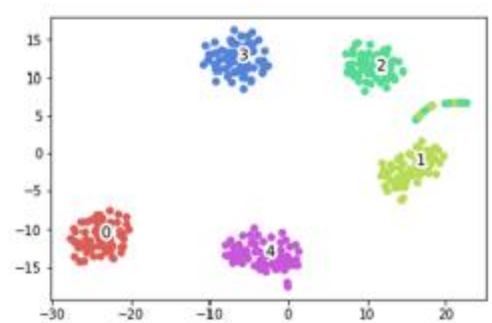
(a) encoder data of standard Auto-encoder by unsupervised learning



(b) encoder data of the proposed model by unsupervised learning



(c) encoder data of the proposed model by supervised learning



(d) frequency-domain data transform from FFT

Feature visualization of SEU gearbox dataset

(a) is the feature that learned by a Standard auto-encoder by unsupervised learning process.

(b) is the feature that learned by the proposed model by unsupervised learning process.

(c) is the feature that learned by the proposed model by supervised learning process.

(d) the frequency-domain signals.

Performance in the same working conditions

- CWRU bearing dataset

	Subsets				
	A	B	C	D	E
Training Set	1hp	2hp	3hp	1-3hp	1-2hp
Testing Set	1hp	2hp	3hp	1-3hp	3hp

Result of the proposed model and compare it with existing methods

Fault diagnosis methods		Dataset					Feature size	Feature extraction
		A	B	C	D	E		
ISA Trans. 2020	DAE				71.26%		1024	No
	MLP				70.11%		1024	No
	CNN				99.62%		1024	No
Sensors 2017	ResNet18				100%		1024	No
	WKDCNN				100%		1024	No
TII, 2018	VGG-16	99.30%	90.66%	99.72%	98.85%	96.47%	1024	Yes
	VGG-16TL	100%	100%	99.96%	99.95%	98.80%	1024	Yes
WKDCAE		100%	100%	100%	99.92%	99.42%	1024	No
RWKDCAE		100%	100%	100%	100%	99.85%	1024	No

Performance in the same working conditions

- SEU gearbox dataset

Result of the proposed model and compare it with existing methods

Fault diagnosis methods		Bearing			Gear			Mixture
		20-0	30-2	All	20-0	30-2	All	All
	SAE-DNN	87.50%	92.10%		92.70%	91.90%		
[9]	GRU	91.20%	92.40%		93.80%	90.50%		
	BiGRU	93.00%	93.60%		93.80%	90.70%		
[6]	LFGRU	93.20%	94.00%		94.80%	95.80%		
	VGG-16TL	99.94%	99.42%		99.64%	99.02%		
[8]	Resnet18							99.50%
WKDCAE		97.38%	97.38%	97.05%	100%	100%	100%	99.51%
RWKDCAE		100%	100%	100%	100%	100%	100%	99.67%

Performance in the different working conditions

- CWRU bearing dataset

Fault diagnosis methods		A → B	A → C	B → A	B → C	C → A	C → B	Average
Zhang et al. (2017)	FFT-SVM	68.60%	60.00%	73.2%	67.6%	68.4%	62.0%	66.6%
	FFT-MLP	82.10%	85.6%	71.5%	82.4%	81.8%	79.0%	80.4%
	WKDCNN	99.40%	93.4%	97.5%	97.2%	88.3%	99.9%	95.9%
WKDCAE		100%	91.7%	99.0%	100%	90.0%	99.0%	96.6%
RWKDCAE		100%	95.0%	99.3%	100%	93.0%	100%	97.9%

- SEU gearbox dataset

Fault diagnosis methods	Bearing 20 → 30	Bearing 30 → 20	Gear 20 → 30	Gear 30 → 20	Average
WKDCAE	78.43%	82.07%	81.51%	76.75%	79.69%
RWKDCAE	83.75%	80.67%	81.67%	77.19%	80.82%

Performance in the noise working conditions

- CWRU bearing dataset

Load	Model	SNR (dB)					
		0	2	4	6	8	10
A	1-DCNN	96.00%	98.33%	99.33%	99.33%	99.33%	99.33%
	1-D WDCNN	96.67%	99.00%	99.00%	99.67%	99.67%	99.67%
	RWKDCAE	99.00%	99.33%	99.67%	99.67%	100%	100%
B	1-DCNN	97.00%	99.33%	99.67%	99.67%	100%	100%
	1-D WDCNN	97.67%	99.33%	99.67%	99.67%	100%	100%
	RWKDCAE	99.33%	100%	100%	100%	100%	100%
C	1-DCNN	96.67%	98.00%	99.33%	100%	100%	100%
	1-D WDCNN	97.67%	98.00%	99.67%	100%	100%	100%
	RWKDCAE	98.00%	100%	100%	100%	100%	100%
D	1-DCNN	97.83%	98.33%	98.33%	98.67%	99.33%	99.33%
	1-D WDCNN	98.25%	98.75%	98.75%	98.92%	99.42%	99.42%
	RWKDCAE	99.08%	99.17%	99.42%	99.58%	99.67%	99.83%

Performance in the noise working conditions

- SEU gearbox dataset

Load	Model	SNR (dB)					
		0	2	4	6	8	10
Bearing 20	1-DCNN	60.46%	90.20%	91.50%	94.77%	96.73%	96.73%
	1-D WDCNN	94.77%	94.77%	94.77%	96.73%	97.39%	97.39%
	RWKDCAE	95.42%	95.42%	96.73%	96.73%	98.69%	98.69%
Bearing 30	1-DCNN	93.46%	94.77%	96.08%	96.08%	96.73%	97.39%
	1-D WDCNN	94.77%	96.08%	96.73%	98.04%	98.04%	99.35%
	RWKDCAE	97.39%	97.39%	98.04%	98.04%	98.69%	99.35%
Gear 20	1-DCNN	81.70%	84.97%	84.97%	84.97%	86.93%	88.89%
	1-D WDCNN	86.27%	90.85%	92.81%	96.08%	98.04%	99.35%
	RWKDCAE	89.54%	94.12%	96.73%	97.39%	98.04%	100%
Gear 30	1-DCNN	68.23%	73.20%	81.70%	86.27%	86.27%	88.89%
	1-D WDCNN	77.12%	88.24%	89.54%	94.12%	97.39%	98.69%
	RWKDCAE	78.43%	90.85%	98.04%	100%	100%	100%
All	1-DCNN	73.20%	80.39%	83.98%	85.45%	86.76%	88.89%
	1-D WDCNN	82.52%	85.78%	86.43%	91.18%	92.97%	94.44%
	RWKDCAE	85.12%	88.89%	92.81%	96.40%	96.40%	97.06%

Performance with different training proportions

- CWRU bearing dataset

Proportion	Models	Datasets			
		A	B	C	D
10%	1-DCNN	88.00%	87.57%	94.78%	97.44%
	1-D WDCNN	89.67%	88.78%	95.44%	98.05%
	RWKDCAE	90.22%	92.78%	98.33%	99.53%
30%	1-DCNN	97.57%	98.00%	99.00%	99.21%
	1-D WDCNN	98.57%	99.00%	99.86%	99.36%
	RWKDCAE	98.86%	99.86%	100%	99.68%
50%	1-DCNN	98.80%	99.40%	100%	99.75%
	1-D WDCNN	98.80%	99.40%	100%	99.76%
	RWKDCAE	99.80%	100%	100%	99.90%

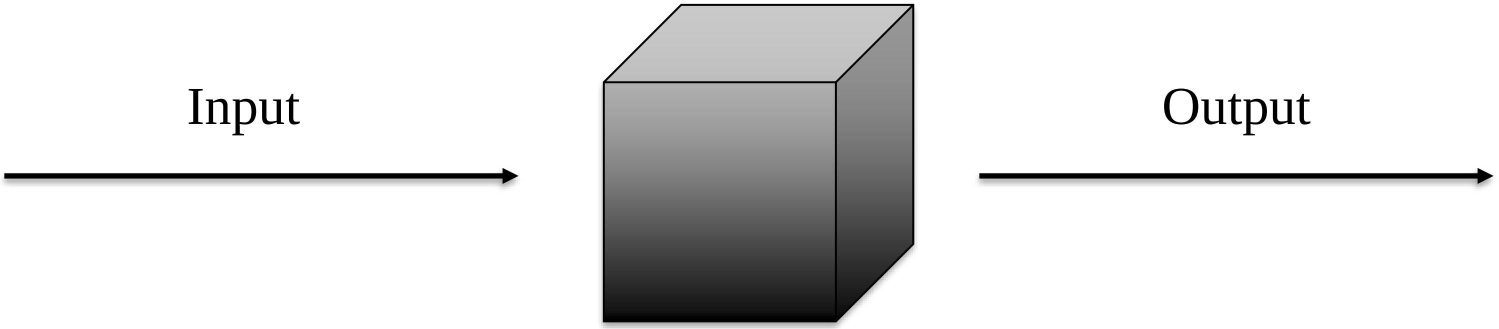
Outline

04 CNN-based explainable fault diagnosis

05 Multi-source information fusion

Basic theory (CNN and explanation methods for CNN)

- ❑ Artificial intelligence algorithms have powerful nonlinear capabilities to deal with a variety of domain problems. However, there is a very serious drawback of many AI algorithms, the Black Box.
- ❑ People wonder if the AI model is making this right decision based on the important parts, instead of the noise parts, this is a question.



Artificial intelligence model

Explainable AI

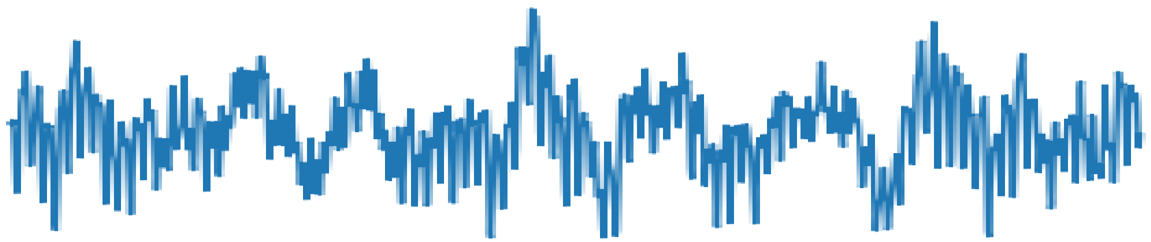
- To understand what the black box model learned in the fault diagnosis framework.

- Develop an explainable intelligence fault diagnosis framework based on post-hoc visualization methods.

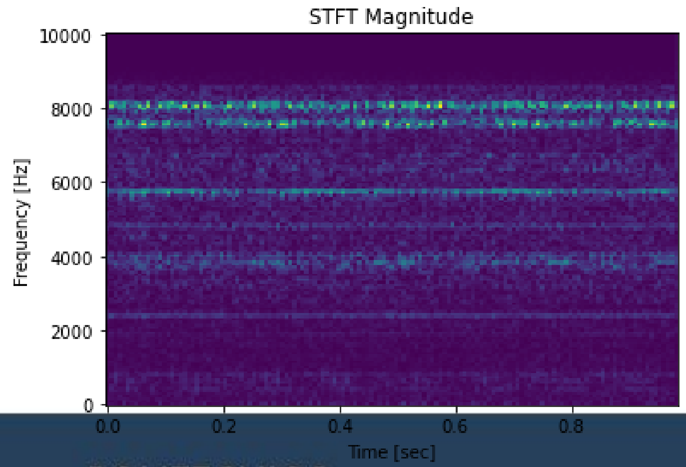
- Compare the performance of post-hoc visualization methods.

Basic theory (CNN and explanation methods for CNN)

- Convolutional neural networks (CNN) is one of the most popular deep learning algorithms.



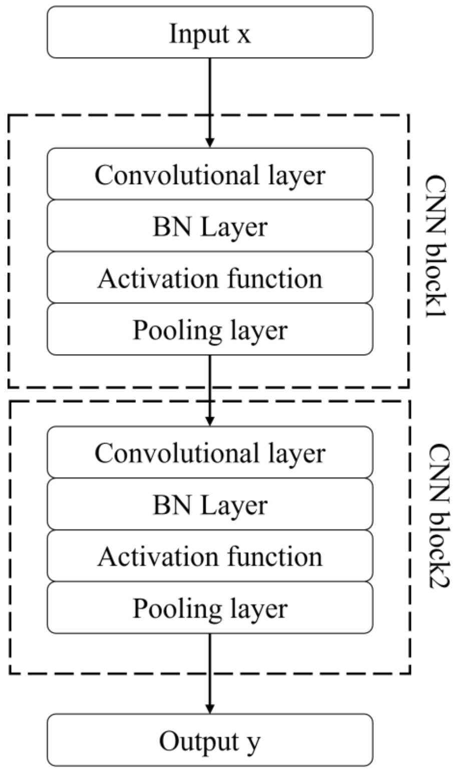
The shape of the time-domain vibration signal



The shape of the time-frequency spectrum

Basic theory (CNN and explanation methods for CNN)

CNN usually consists of four different layers in the convolutional block, including convolutional layer, batch-normalization layer, activation layer and pooling layer.



$$a = W \otimes x + b$$

$$s = BN(a)$$

$$h = ReLU(s) = \max(0, s)$$

$$y = MaxPool(h)$$

Basic theory (CNN and explanation methods for CNN)
Explanation methods for CNN

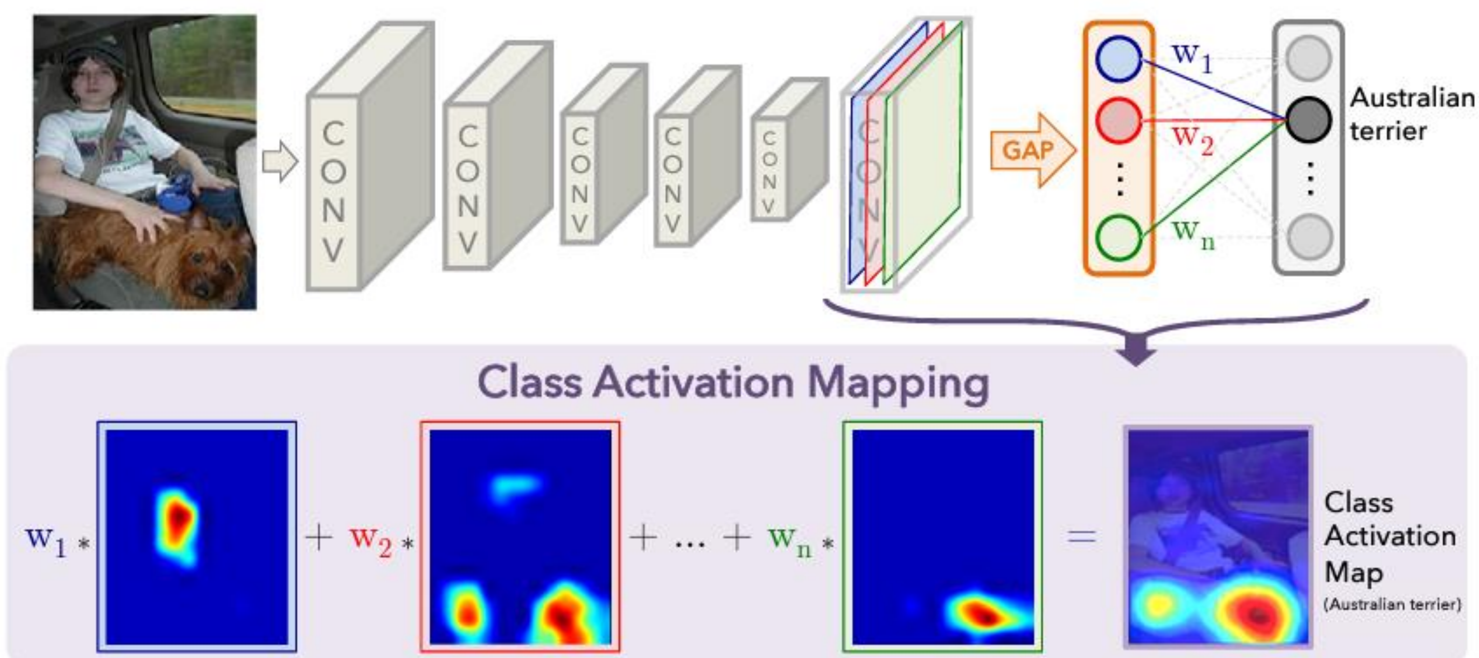
- Convolutional neural networks (CNN):
  - Advantages:
    - Strong capability.
    - Without expert rotating machinery fault diagnosis knowledge.
  - Disadvantage:
    - It is a black box model.

- Post-hoc visualization methods

1) Classification activation map (CAM)
2) Gradient-weighted classification activation map (Grad-CAM)
3) Gradient-weighted classification activation map ++ (Grad-CAM++)
4) Score classification activation map (Score CAM)

Classification activation mappings (CAM)

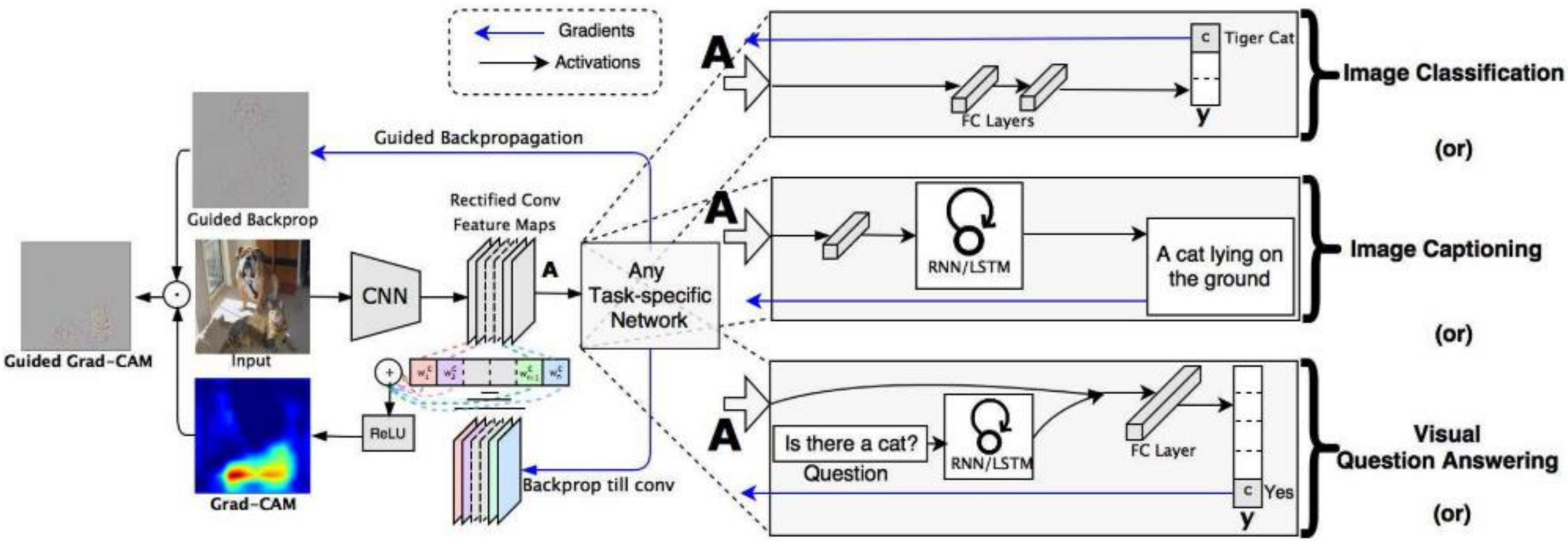
- It is a simple method to get the saliency maps of the CNN models.
- But it needs to change the structure of the CNN model (Global average pooling layer).



Schematic diagram of CAM.

Grad-CAM

- It is no need to change the structure of the CNN model.
  - It could be used in many kinds of tasks.
  - Coarse localization maps



Grad-CAM++

- Better localization performance than Grad-CAM.
- Better visualization performance when there are several features in one input data than Grad-CAM.

Grad-CAM:

$$Y^c = \sum_k w_k^c \cdot \sum_i \sum_j A_{ij}^k$$

$$w_k^c = Z \cdot \frac{\partial Y^c}{\partial A_{ij}^k}, \ \forall \{i,j | i,j \in A^k\}$$

$$L_{Grad-CAM}^c = ReLU\left(\sum_k w_k^c \cdot A_{ij}^k\right)$$

* Z is a constant

Grad-CAM++:

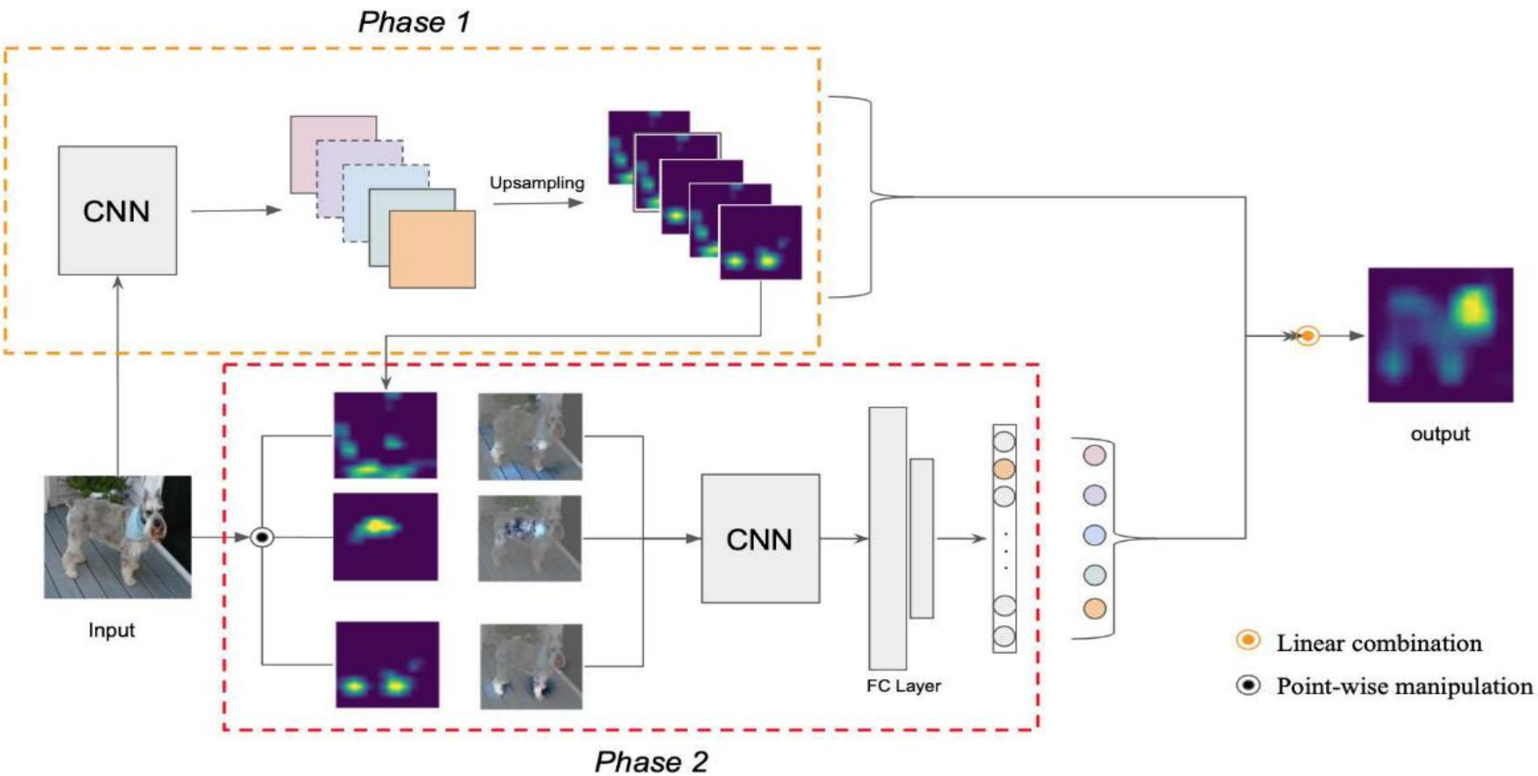
$$w_k^c = \sum_i \sum_j \alpha_{ij}^{kc} \cdot ReLU\left(\frac{\partial Y^c}{\partial A_{ij}^k}\right)$$

$$\alpha_{ij}^{kc} = \frac{\dfrac{\partial^2 Y^c}{\left(\partial A_{ij}^k\right)^2}}{2 \cdot \dfrac{\partial^2 Y^c}{\left(\partial A_{ij}^k\right)^2} + \sum_a \sum_b A_{ab}^k \left\{\dfrac{\partial^3 Y^c}{\left(\partial A_{ij}^k\right)^3}\right\}}$$

$$L_{Grad-CAM++}^c = ReLU\left(\sum_k w_k^c \cdot A_{ij}^k\right)$$

Score-CAM

- Better localization precise.
- Better visualization performance without gradient.

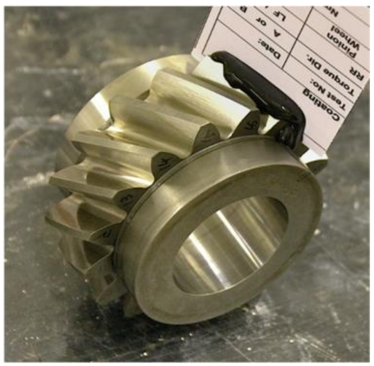


Schematic diagram of Score-CAM.

Experiment

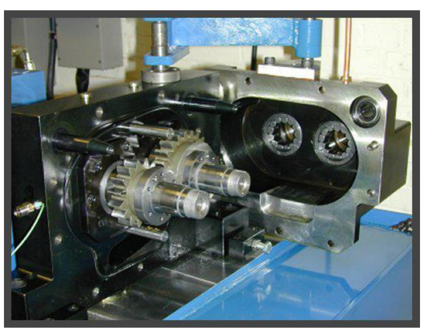
Gearbox dataset

- Test gears in modules from 2 to 6 mm, face widths up to 30 mm, helix angles up to 30^{o}, torques up to 1400 N·m.
- 16-teeth pinion, 24-teeth wheel.



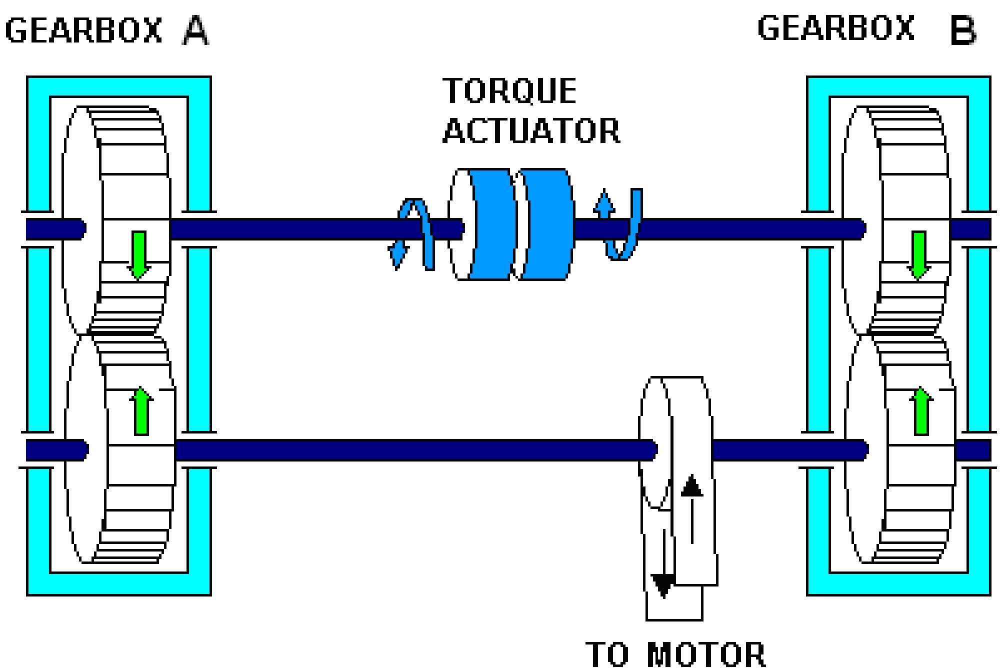
16-teeth pinion

24-teeth wheel



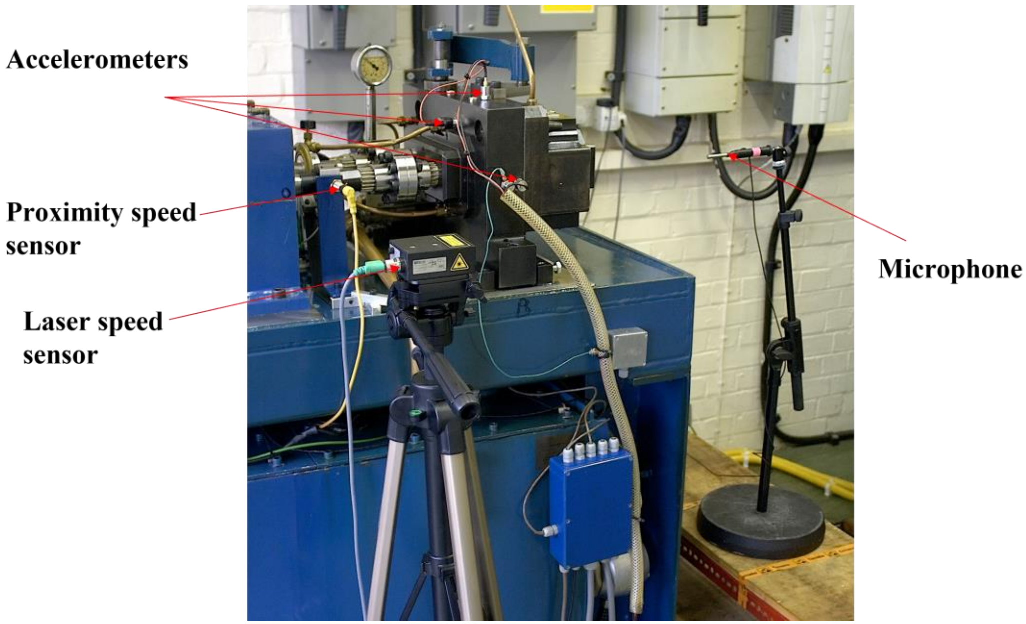
Schematic diagram and picture of the gearbox test rig.

Experiment
Gearbox dataset



- The rig is a power-recirculating rig, with two identical test gearboxes A and B connected via torsionally compliant shafts.
- A servo-hydraulic torque actuator is interposed between the gearboxes, allowing precise closed-loop control of torque, and adjustment whilst running.

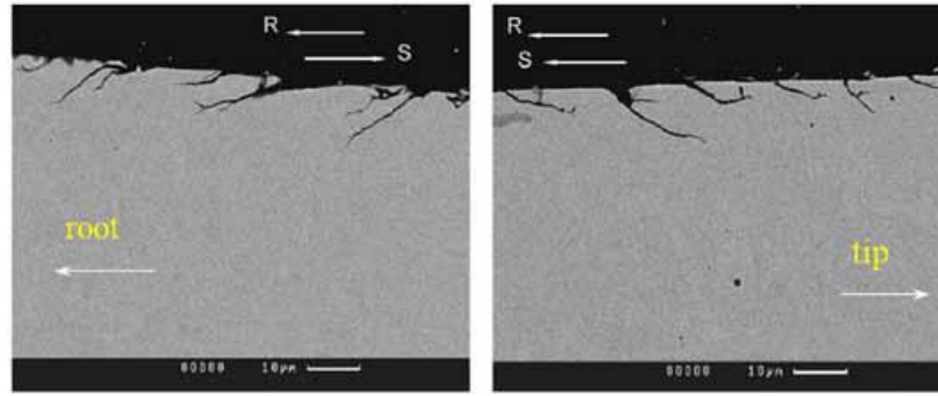
Experiment
Gearbox dataset



- This actual test is a constant Torque was performed.
- At the first part with wheel torque set at 500 +/- 5 Nm and 50 million cycles (shaft revolutions) was performed
- After completion of each 10 million cycles (2.5 days approximately) test has been stopping for gears assessment and then ran again

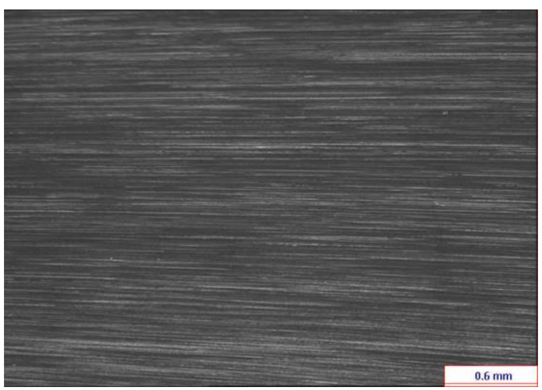
Experiment
Gearbox dataset

- Understanding the morphology of micropitting is the key to determining the primary failure mode and root cause of failure.
- Micropitting cracks grow opposite the direction of sliding at the gear tooth surface.

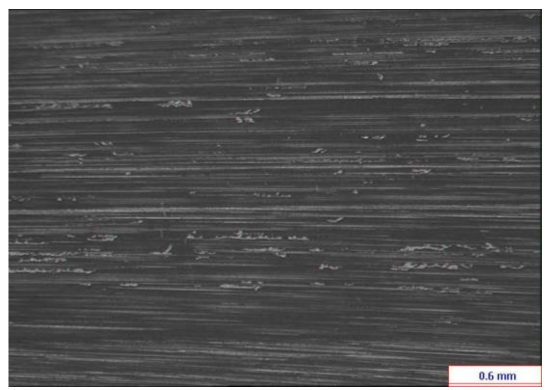


Micropitting cracks on a driven gear (courtesy of Newcastle University)

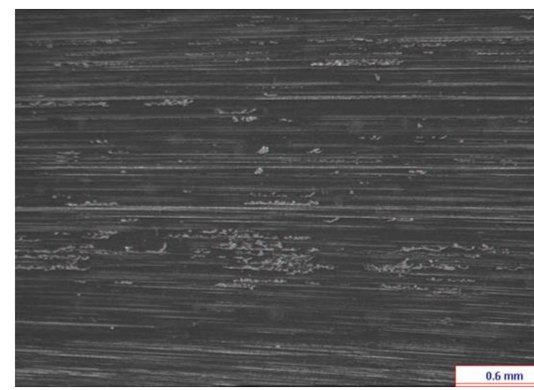
Progress of micro-pitting for tooth 1



As ground



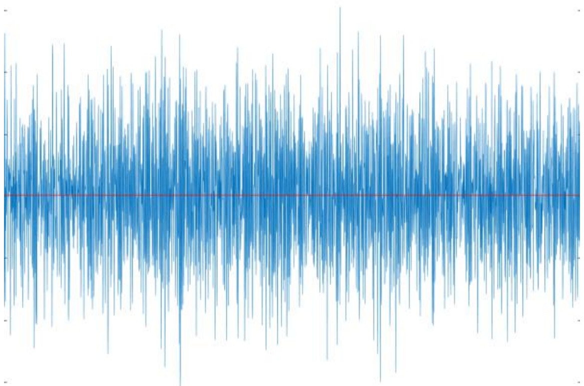
After 10 million cycles



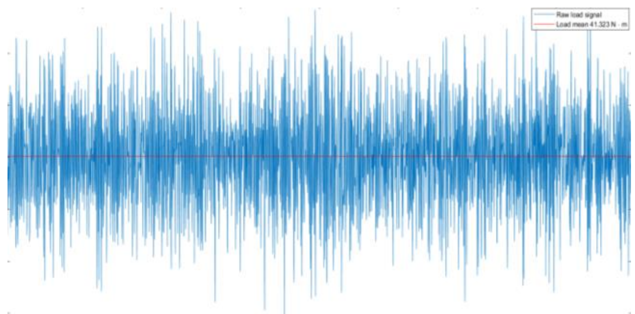
After 30 million cycles

Experiment
Gearbox dataset

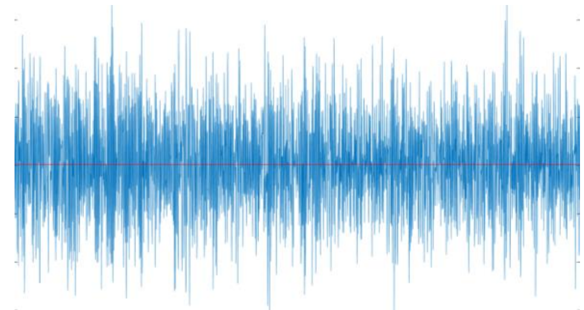
Vibration signals and Frequency-domain signals



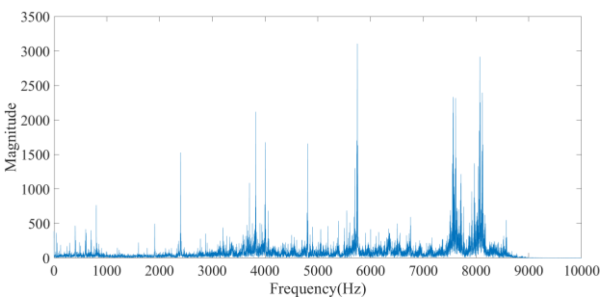
As ground

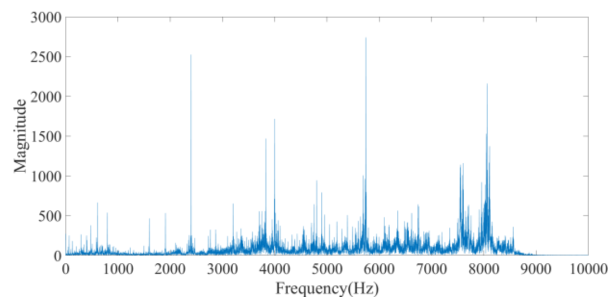


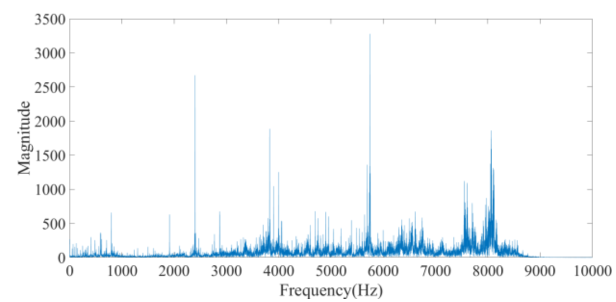
After 10 million cycles



After 30 million cycles

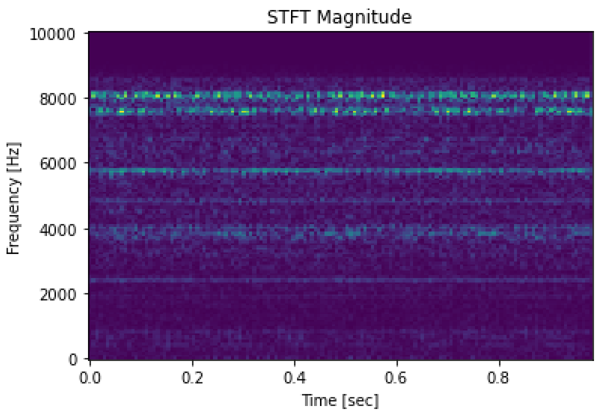




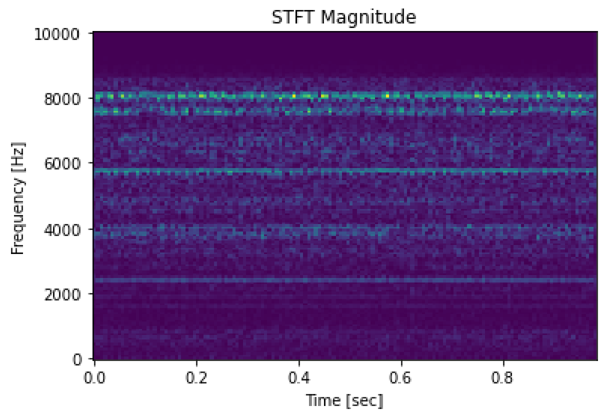


Experiment
Gearbox dataset

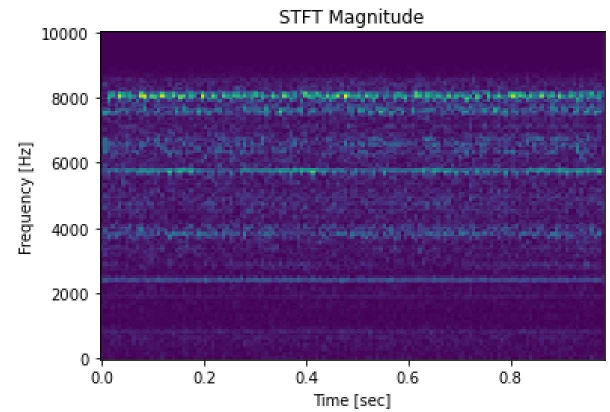
Time-frequency domain signals (STFT)



As ground

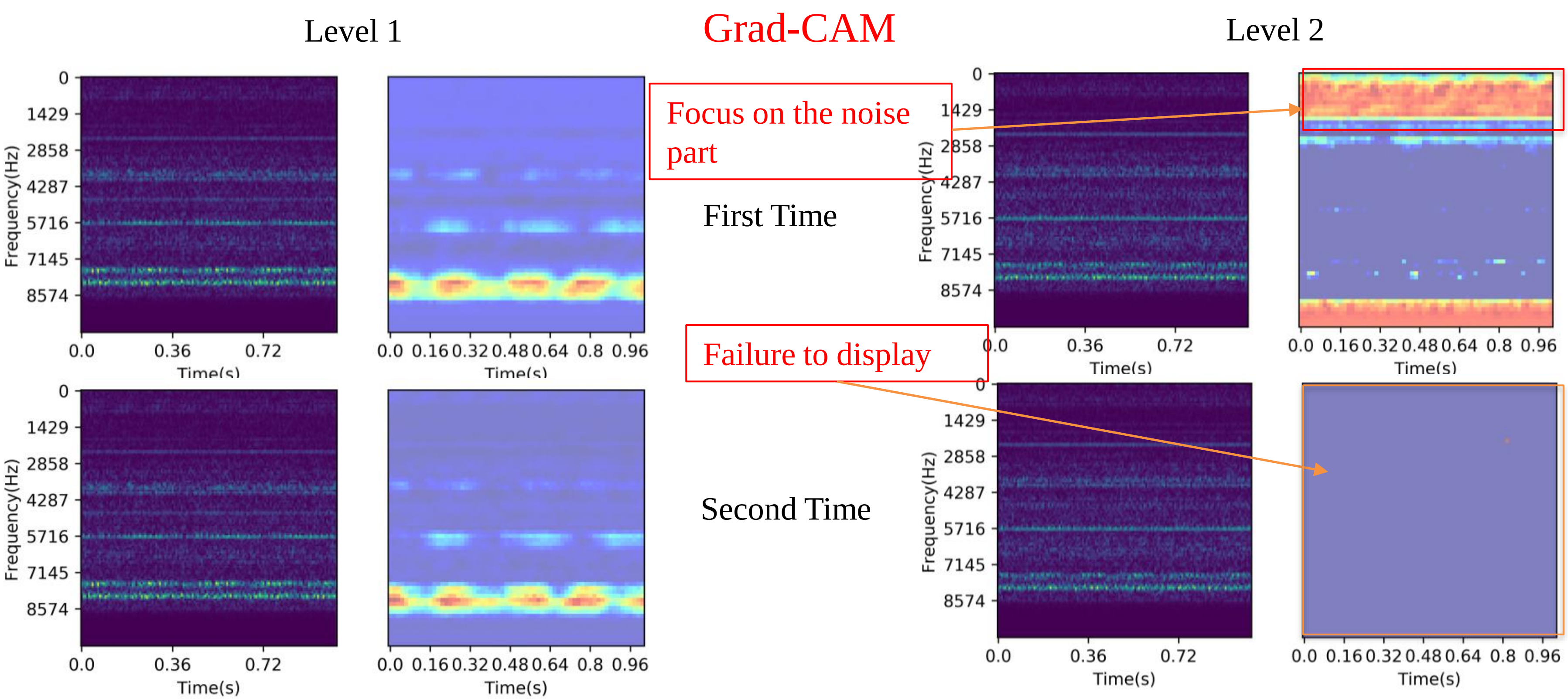


After 10 million cycles



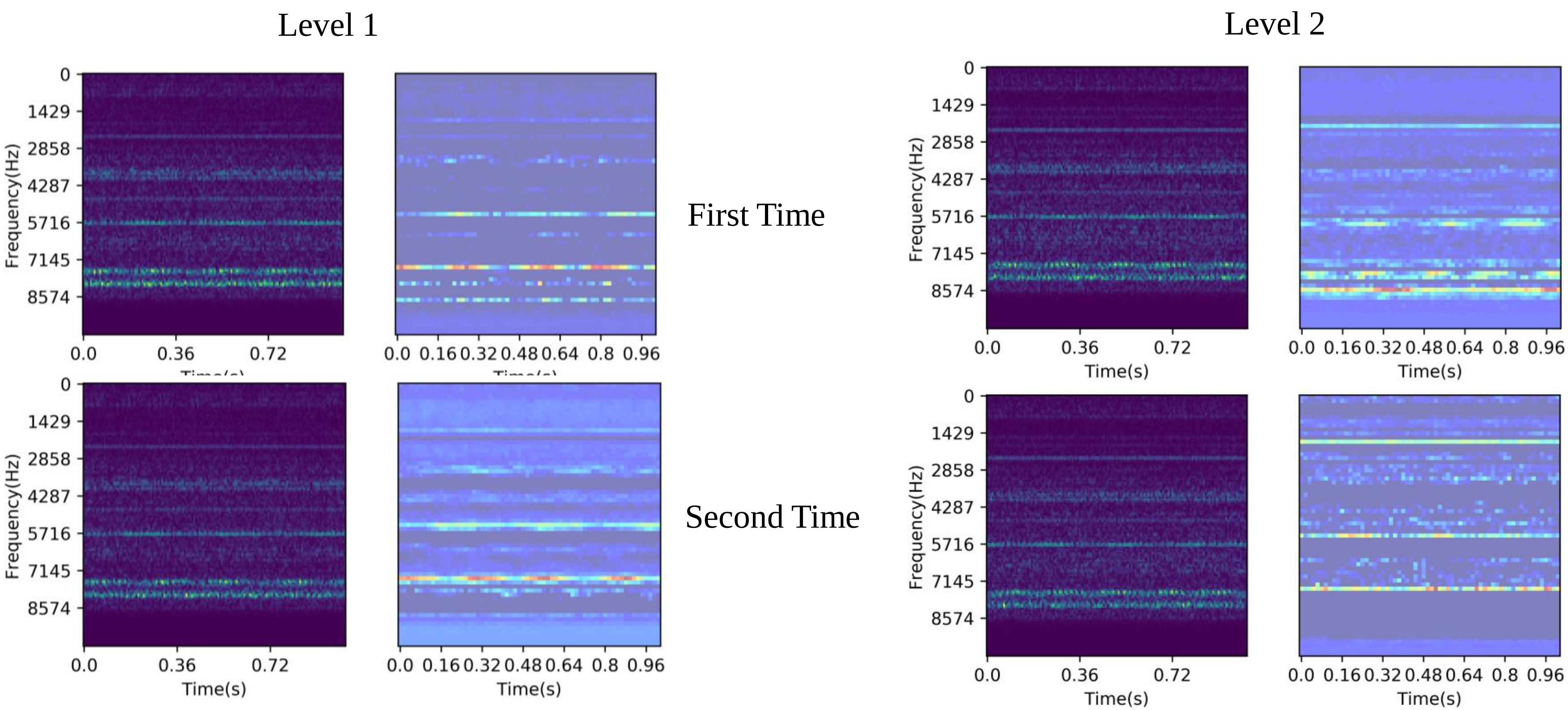
After 30 million cycles

Comparative results of Grad-CAM, Grad-CAM++ and Score-CAM



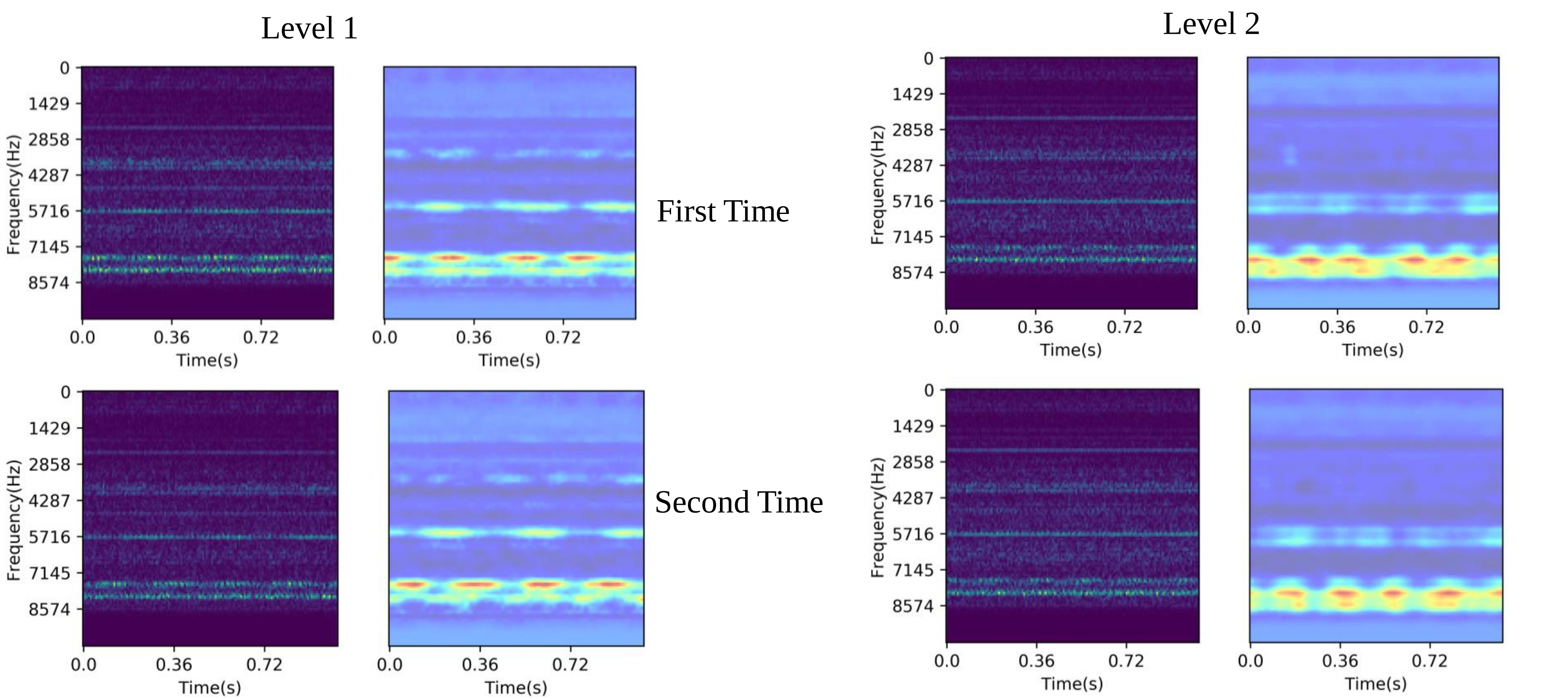
Comparative results of Grad-CAM, Grad-CAM++ and Score-CAM

Grad-CAM++



Comparative results of Grad-CAM, Grad-CAM++ and Score-CAM

Score-CAM



Comparative results of Grad-CAM, Grad-CAM++ and Score-CAM

Average Drop:

The original input is masked by pointwise multiplication with the saliency maps to observe the score change on the target class.

$$Average\ Drop = \sum_{i=1}^{N} \frac{max(0, Y_i^c - O_i^c)}{Y_i^c} \times 100\%$$

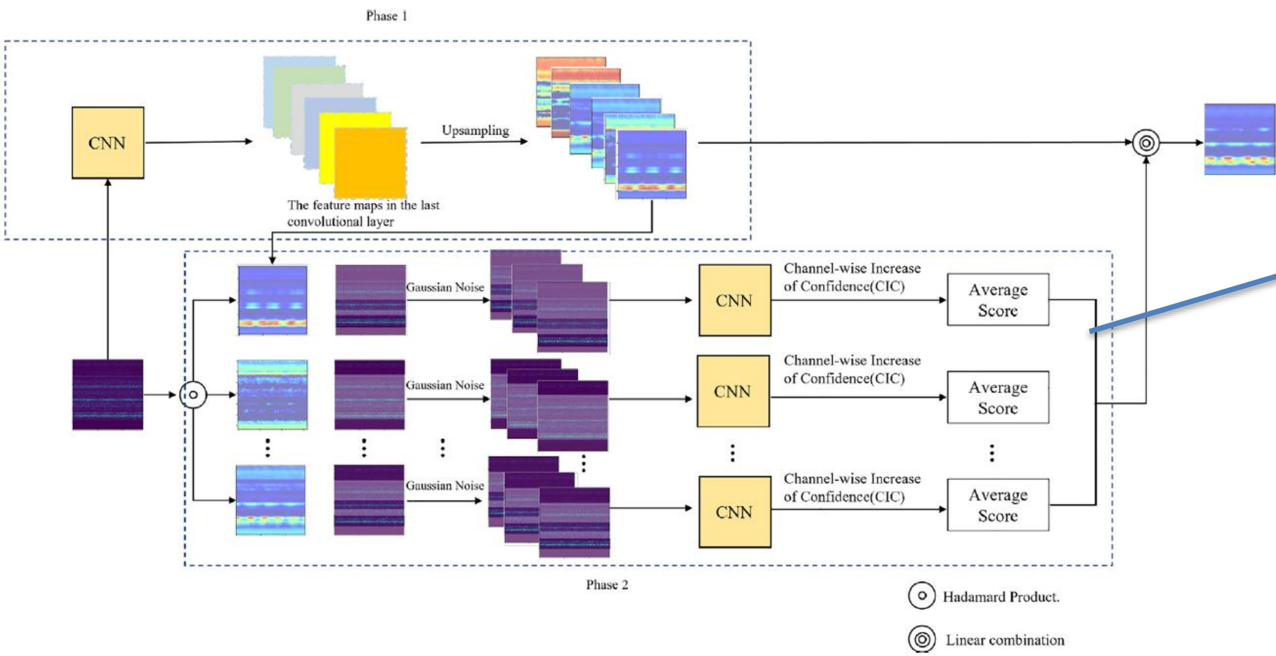
Y_i^c is the score with the raw input data

O_i^c is the score with the new input by the pointwise multiplication with the saliency maps

Table 1, the Average Drop(%) of Grad-CAM++, Grad-CAM and Score-CAM

Method	Grad-CAM++	Grad-CAM	Score-CAM
Average Drop(%) (Lower is better)	79.59	83.47	40.88

Smoothed Score-CAM(SS-CAM)



$$C\left(A_l^k\right) = f\left(X \circ H_l^k\right) - f(X_b)$$

$$H_l^k = s\left(U\left(A_l^k\right)\right)$$

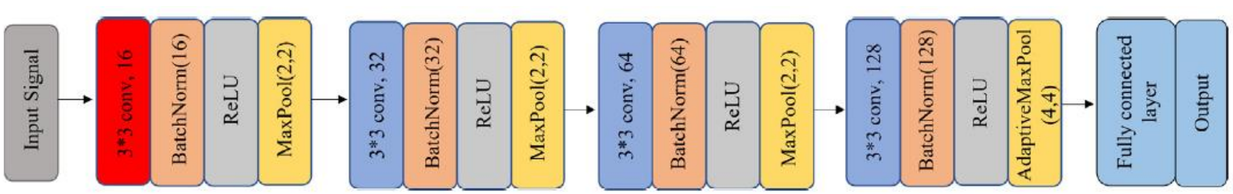
$$s\left(A_l^k\right) = \frac{A_l^k - min\left(A_l^k\right)}{max\left(A_l^k\right) - min\left(A_l^k\right)}$$

$$\eta_c^k = \frac{\sum_1^N (C(M))}{N}$$

$$M = \sum_1^N \left(X * \left(A_l^k + N(0, \sigma^2)\right)\right)$$

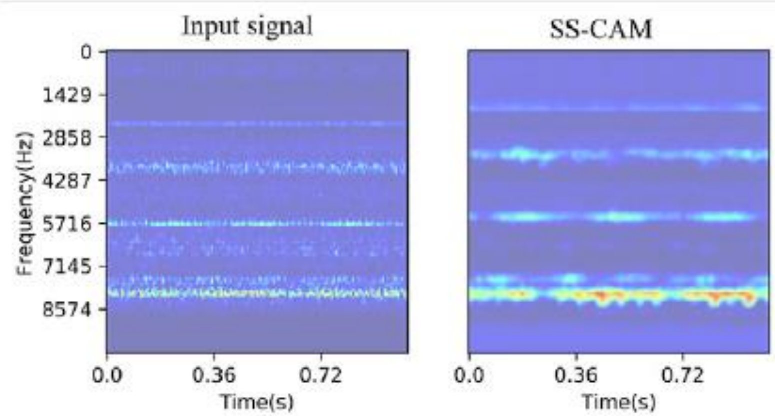
$$L_{SS-CAM}^c = ReLU\left(\sum_k \eta_k^c A_l^k\right)$$

The structure of the standard CNNs.

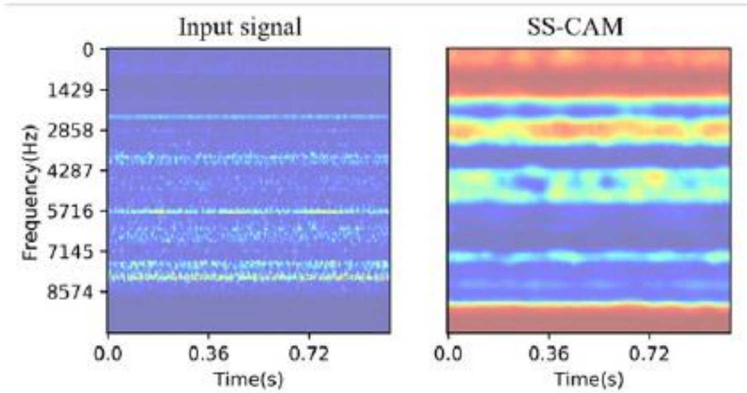


Smoothed Score-CAM(SS-CAM)

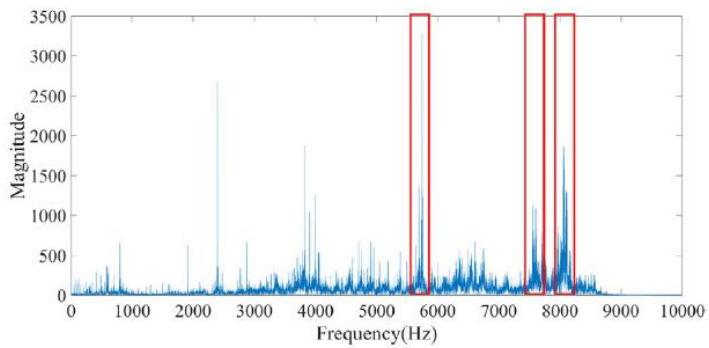
Localization evaluation results for Fault severity level 4

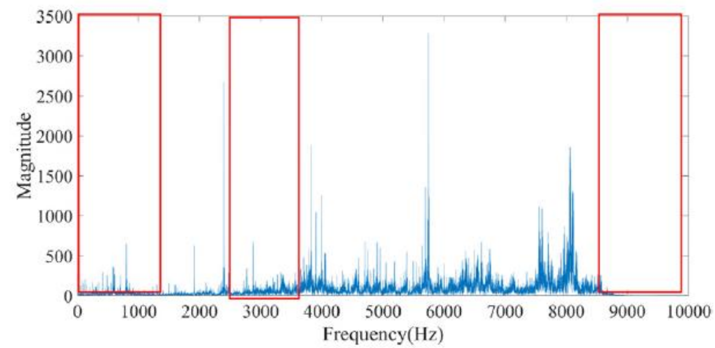


The proposed intelligence fault diagnosis method.



The standard CNNs





Outline

Multi-source open set domain adaptation for fault diagnosis

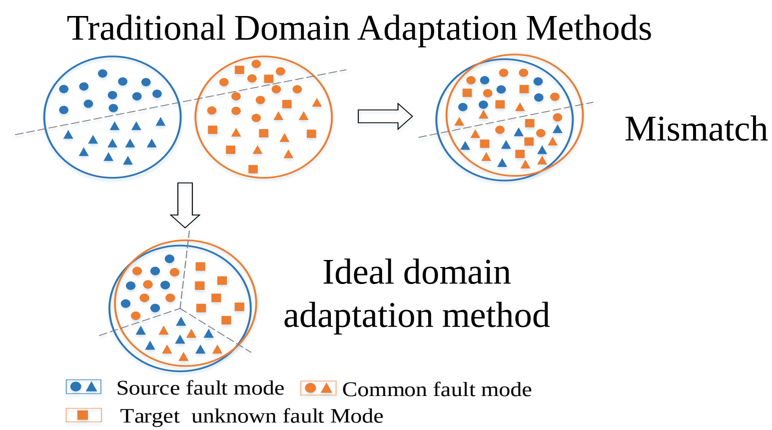
Question: Mechanical equipment is a multi-level and nonlinear complex whole, which is easy to induce new faults, so that the diagnostic knowledge of the source domain cannot completely cover the fault categories of the target domain.

Previous work:

- ➢ Traditional domain adaptation methods that align marginal distributions without isolating unknown fault instances can lead to model failure due to mismatches between known and unknown classes.

Application scenario:

- ✓ The data distribution of multiple source domains is different from that of the target domain, and the number of fault types in the source domain is smaller than that in the target domain;
- ✓ The target domain samples are unlabeled.



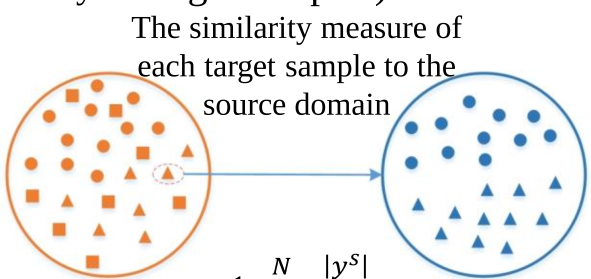
Challenge:

- ➢ how to adapt the characteristics of the shared classes in the two domains, and realize the detection of known faults and new faults at the same time.

Compared with closed-set domain adaptation, open-set domain adaptation allows unseen classes to exist in the test scene, and is suitable for unknown class identification.

Multi-source open set domain adaptation for fault diagnosis

Similarity metrics (combining multiple complementarity indicators predicted by multiple source classifiers to measure the transferability of target samples)



$$w_i^{\mathrm{ent}}(\hat{y}_i^{tj}|_{j=1}^{N}) = \frac{1}{N}\sum_{j=1}^{N}\Big(\sum_{k=1}^{|y^s|} -\hat{y}_{ik}^{tj}\log(\hat{y}_{ik}^{tj})\Big)$$

$$w_i^{\mathrm{conf}}(\hat{y}_i^{tj}|_{j=1}^{N}) = \frac{1}{N}\sum_{j=1}^{N}\max(\hat{y}_i^{tj})$$

$$w_i^{\mathrm{cons}}\left(\hat{y}_i^{tj}|_{j=1}^{N}\right) = \frac{1}{|y^s|}\left\|\frac{1}{N}\sum_{j=1}^{N}\left(\hat{y}_i^{tj} - \frac{1}{N}\sum_{j=1}^{N}\hat{y}_i^{tj}\right)^2\right\|_1$$

$$w_i^t = \frac{(1-w_i^{\mathrm{ent}}) + (1-w_i^{\mathrm{cons}}) + w_i^{\mathrm{conf}}}{3}$$

Weighted adversarial mechanism (eliminate the influence of outlier sample points in distribution matching by weighting target samples in adversarial training)

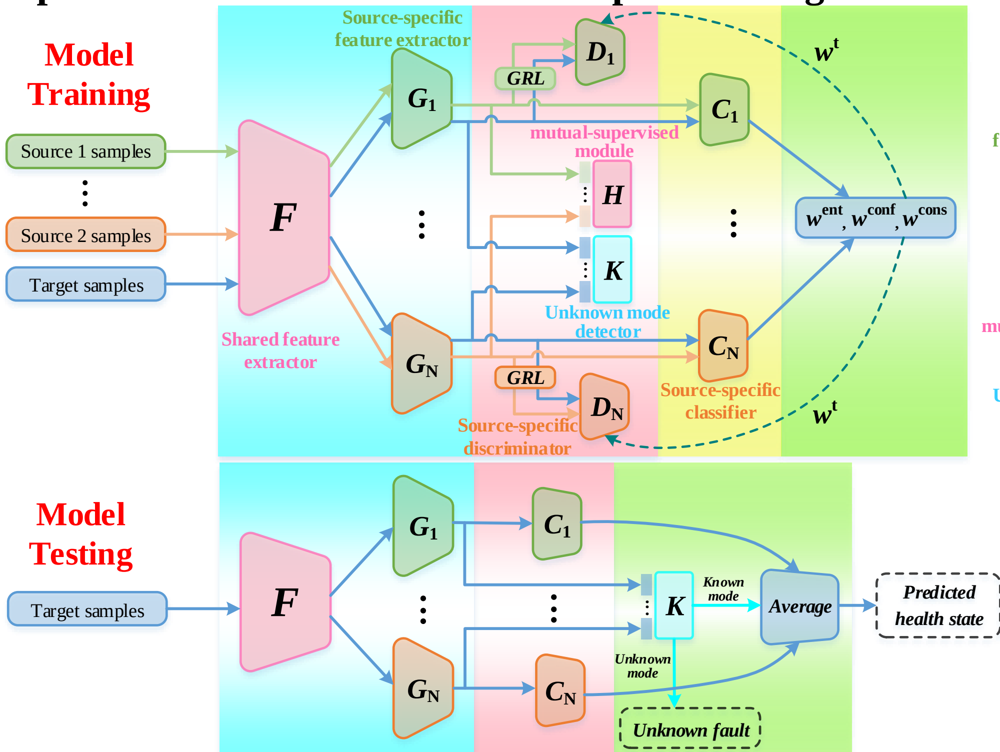
$$L_{\mathrm{adv}}\left(\theta_F, \theta_{G_j}, \theta_{D_j}\right) = \frac{1}{N_j}\sum_{i=1}^{N_j} J\left(D_j\left(G_j\left(F\left(x_i^{sj}\right)\right)\right), d_i\right) + \frac{1}{N_t}\sum_{i=1}^{N_t} w_i^t\, J(D_j\left(G_j\left(F(x_i^t)\right)\right), d_i)$$

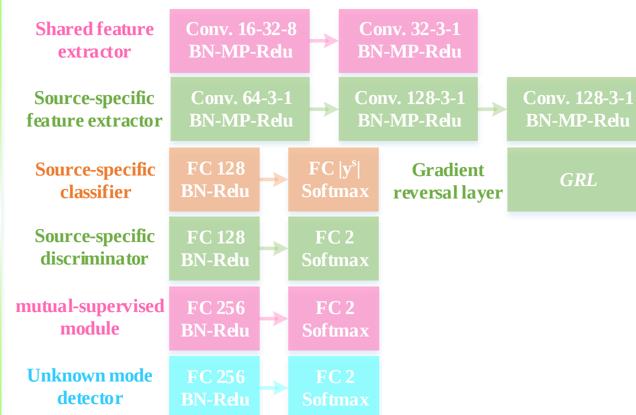
Pseudo-labeled samples to train unknown mode detectors (isolating unknown fault samples by unknown mode detectors)

$$L_{\mathrm{uk}}\left(\theta_F, \theta_{G_j}|_{j=1}^{N}, \theta_K\right) = \frac{1}{N_*}\sum_{i=1}^{N_*} J\left(K\left(G_1\left(F(x_i^t)\right), \cdots, G_N\left(F(x_i^t)\right)\right), k_i\right)$$

Multi-source open set domain adaptation for fault diagnosis

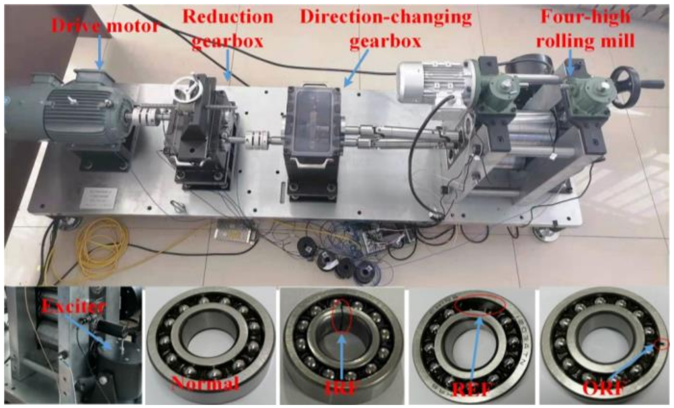
Open-set Multi-source Domain Adaptation Diagnosis Model (OSMDA)



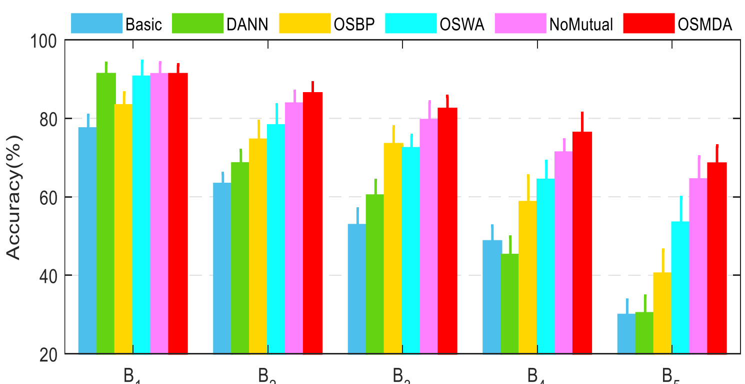


A multi-branch network structure is adopted to synthesize the prediction results of multiple specific source classifiers, and a complementarity metric is applied to quantify the similarity between each target sample and the source domain.

Multi-source open set domain adaptation for fault diagnosis



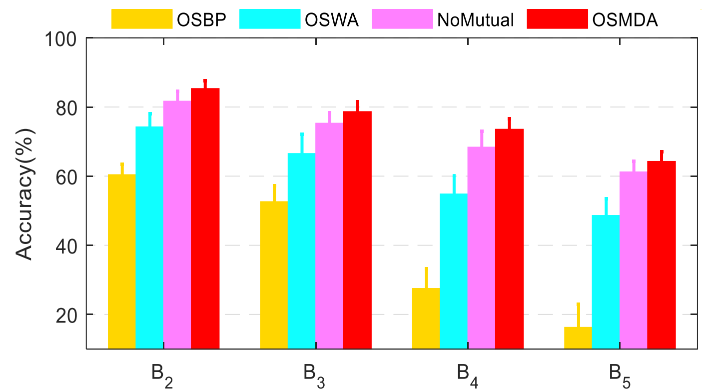
Rolling mill test platform



Comparison of Diagnostic Accuracy of Different Methods

Diagnostic tasks under different degrees of openness

Task	Multi-source to Target	Source states	Openness
B_1	600, 1200→1800	All	0
B_2	600, 1200→1800	0, 1, 3, 4, 5	1/6
B_3	1200,1800→600	0, 1, 3, 5	1/3
B_4	1200,1800→600	0, 2, 3	1/2
B_5	600,1800→1200	0, 1	2/3

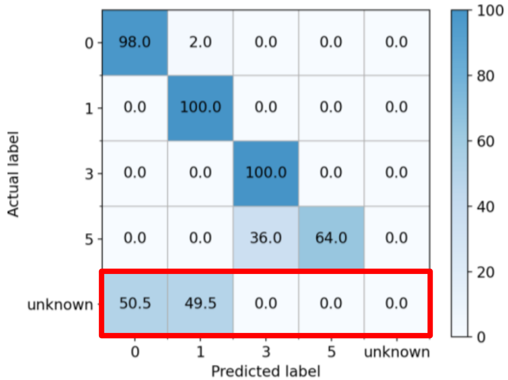


Comparison of detection accuracy of unknown faults by different methods

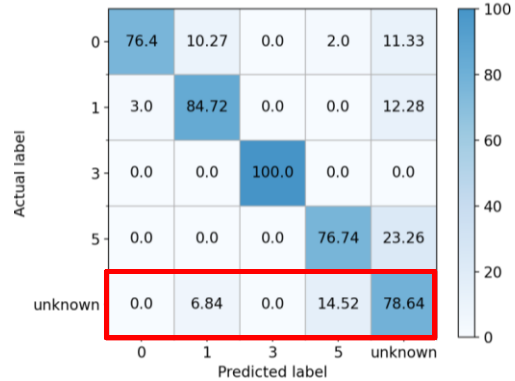
In different diagnostic tasks, the proposed method achieves the highest diagnostic accuracy, which can effectively identify common faults and detect unknown fault modes.

Multi-source open set domain adaptation for fault diagnosis

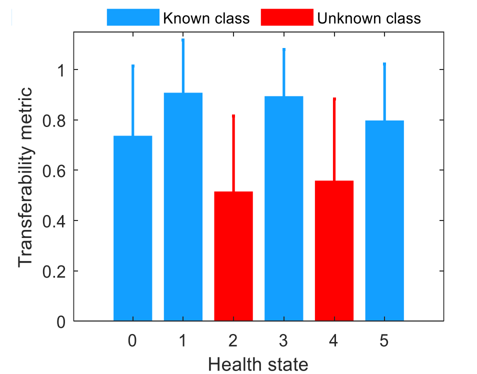
Classification confusion matrix and feature visualization



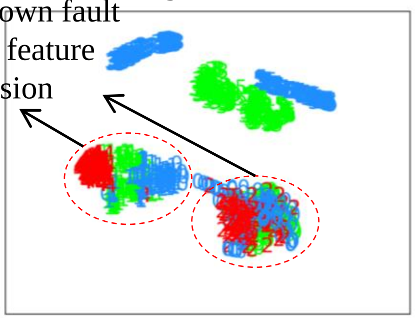
Task B_3 –DANN



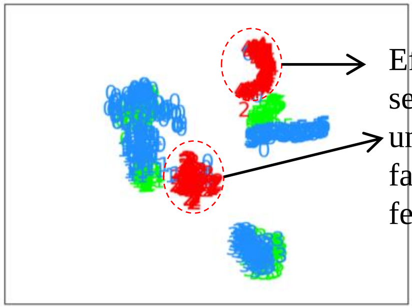
Task B_3 –OSMDA



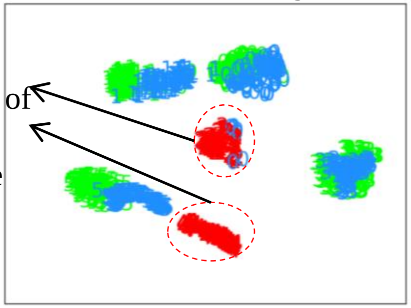
Average similarity weight of different classes under Task B_3



Task B_3 –DANN



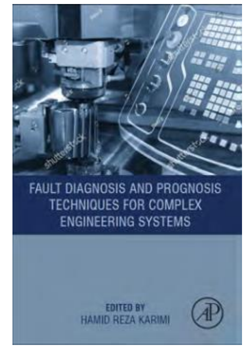
Task B_3 –OSMDA (D_{s1}-D_t)



Task B_3 –OSMDA (D_{s2}-D_t)

References

- Fault Diagnosis and Prognosis Techniques for Complex Engineering Systems, 1st Edition, ISBN: 9780128224731, 2021



- Vibration analysis for bearing fault detection and classification using an intelligent filter, Mechatronics 24 (2), 151-157, 2014
- A review of diagnostics and prognostics of low-speed machinery towards wind turbine farm-level health management, Renewable and Sustainable Energy Reviews 53, 697-708, 2016
- Residual wide-kernel deep convolutional auto-encoder for intelligent rotating machinery fault diagnosis with limited samples, Neural Networks 141, 133-144, 2021
- "An explainable intelligence fault diagnosis framework for rotating machinery." Neurocomputing: 126257, 2023.
- Deep learning-based open set multi-source domain adaptation with complementary transferability metric for mechanical fault diagnosis, Neural Networks 162, 69-82, 2023

POLITECNICO
MILANO 1863

DIPARTIMENTO DI MECCANICA

Thanks for your attention.

mecc

www.mecc.polimi.it

@meccpolimi