Data-Driven Fault Diagnosis: From Sparse Representation To Deep Learning

Xianchao Xiu

Department of Automation

上海大学
Shanghai University

Mathematical Optimization Society, May 12-14, 2023

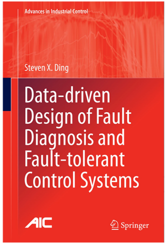
Joint work with Ying Yang (PKU), Wanquan Liu (SYSU) and others

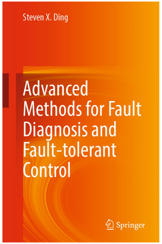
Outline

FD

- Fault diagnosis (FD) is one of the research hotspots in industrial engineering





- Model-based fault diagnosis techniques: design schemes, algorithms, and tools, 2008
- Data-driven design of fault diagnosis and fault-tolerant control systems, 2014
- Advanced methods for fault diagnosis and fault-tolerant control, 2021

PCA

- Principal component analysis (PCA)

$$\begin{aligned} \min_{A} \quad & \frac{1}{2}\|X - AA^\top X\|_F^2 \\ \text{s.t.} \quad & A^\top A = I \end{aligned} \qquad\qquad \begin{aligned} \min_{A} \quad & -\mathrm{Tr}(A^\top X^\top X A) \\ \text{s.t.} \quad & A^\top A = I \end{aligned}$$

- Sparse principal component analysis (SPCA)

$$\begin{aligned} \min_{A} \quad & \frac{1}{2}\|X - AA^\top X\|_F^2 + \lambda\|A\|_{2,1} \\ \text{s.t.} \quad & A^\top A = I \end{aligned} \qquad\qquad \begin{aligned} \min_{A} \quad & -\mathrm{Tr}(A^\top X^\top X A) + \lambda\|A\|_{2,1} \\ \text{s.t.} \quad & A^\top A = I \end{aligned}$$

- Pearson, Philos Mag, 1901
- Zou-Hastie-Tibshirani, Journal of Computational and Graphical Statistics, 2006
- Gewers-Ferreira-Arruda-Silva-Comin-Amancio-Costa, ACM Computing Surveys, 2021
- Greenacre-Groenen-Hastie-Markos-Tuzhilina, Nature Reviews Methods Primers, 2022

PCA

- Liu-Zhang-Xu, JPC, 2017

Contents lists available at ScienceDirect

Journal of Process Control

journal homepage: www.elsevier.com/locate/jprocont

Compressive sparse principal component analysis for process supervisory monitoring and fault detection

Yang Liu[a,*], Guoshan Zhang[b], Bingyin Xu[a]

- Liu-Zeng-Xie-Luo-Su, IEEE TII, 2019

IEEE TRANSACTIONS ON INDUSTRIAL INFORMATICS, VOL. 15, NO. 5, MAY 2019 2721

Structured Joint Sparse Principal Component Analysis for Fault Detection and Isolation

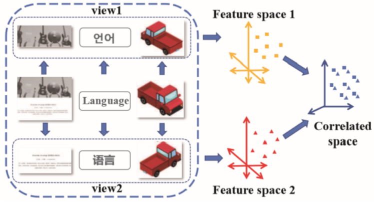
Yi Liu, Jiusun Zeng, Lei Xie, Shihua Luo, and Hongye Su

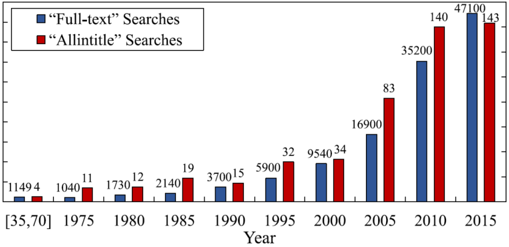
CCA

- Canonical correlation analysis (CCA)

$$\min_{A,B} \quad \frac{1}{2}\|XA - YB\|_F^2$$
$$\text{s.t.} \quad A^\top X^\top XA = I, \; B^\top Y^\top YB = I$$

$$\min_{A,B} \quad -\operatorname{Tr}(A^\top X^\top YB)$$
$$\text{s.t.} \quad A^\top X^\top XA = I, \; B^\top Y^\top YB = I$$





- Hotelling, Biometrika, 1936
- Yang-Liu-Liu-Tao, IEEE TKDE, 2021

CCA

- Statistics
  - Witten-Tibshirani-Hastie, Extensions of sparse canonical correlation analysis with applications to genomic data, Biostatistics, 2009
  - Andrew-Arora-Bilmes-Livescu, Deep canonical correlation analysis, ICML, 2013
  - Lindenbaum-Salhov-Averbuch-Kluger, ℓ_0-sparse canonical correlation analysis, ICML, 2022
- Optimization
  - Chu-Liao-Ng-Zhan, Sparse canonical correlation analysis: New formulation and algorithm, IEEE TPAMI, 2013
  - Chen-Ma-Xue-Zou, An alternating manifold proximal gradient method for sparse principal component analysis and sparse canonical correlation analysis, IJOO, 2020
  - Li-Xiu-Liu-Miao, An efficient Newton-based method for sparse generalized canonical correlation analysis, IEEE SPL, 2022
- Machine Learning
  - Chu-Liao-Ng-Zhan, Sparse canonical correlation analysis: New formulation and algorithm, IEEE TPAMI, 2013
  - Sun-Xiu-Luo-Liu, Learning high-order multi-view representation by new tensor canonical correlation analysis, IEEE TCSVT, 2023
  - Zhou-Ataee-Hou-Tong-X-Feng-Long-Shen, Fair canonical correlation analysis, NeurIPS, 2024

CCA

- Chen-Ding-Zhang-Li-Hu, CEP, 2016

Contents lists available at ScienceDirect

Control Engineering Practice

journal homepage: www.elsevier.com/locate/conengprac

Canonical correlation analysis-based fault detection methods with application to alumina evaporation process ☆

Zhiwen Chen [a,*], Steven X. Ding [a], Kai Zhang [a], Zhebin Li [b], Zhikun Hu [b]

- Chen-Ding-Peng-Yang-Gui, IEEE TIE, 2018

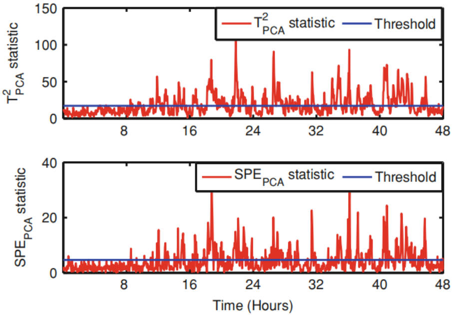
IEEE TRANSACTIONS ON INDUSTRIAL ELECTRONICS, VOL. 65, NO. 2, FEBRUARY 2018 1559

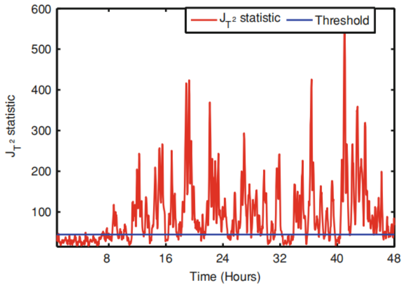
Fault Detection for Non-Gaussian Processes Using Generalized Canonical Correlation Analysis and Randomized Algorithms

Zhiwen Chen, Steven X. Ding, Tao Peng, Chunhua Yang, *Member, IEEE*, and Weihua Gui, *Member, IEEE*

Motivation

- PCA *v.s.* CCA





- What shall we do
  - How to improve performance?
  - How to develop efficient algorithms?
  - How to apply to industrial engineering?

Outline

Robust PCA

- Xiu-Yang-Kong-Liu, JPC, 2020

$$\begin{aligned} \min_{A,B} \quad & \frac{1}{2}\|X - XBA^\top\|_F^2 + \lambda_1\|B\|_{2,1} \\ \text{s.t.} \quad & A^\top A = I \end{aligned}$$

$$\Downarrow$$

$$\begin{aligned} \min_{A,B,E} \quad & \frac{1}{2}\|X - XBA^\top - E\|_F^2 + \lambda_1\|B\|_{2,1} + \lambda_2\|E\|_1 + \lambda_3 \mathrm{Tr}(B^\top L^h B) \\ \text{s.t.} \quad & A^\top A = I \end{aligned}$$

- Alternating direction method of multipliers (ADMM)

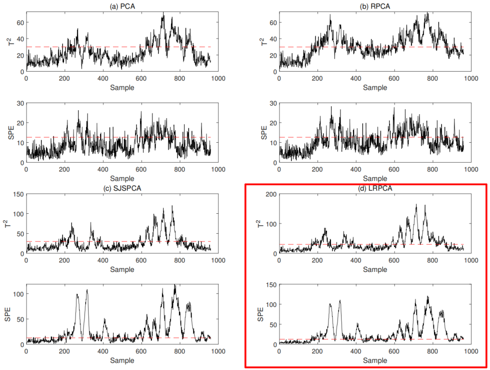
$$\begin{aligned} \min_{A,C,D,E,B} \quad & \frac{1}{2}\|X - XCA^\top - E\|_F^2 + \lambda_1\|D\|_{2,1} + \lambda_2\|E\|_1 + \lambda_3 \mathrm{Tr}(B^\top L^h B) \\ \text{s.t.} \quad & A^\top A = I,\ B = C,\ B = D \end{aligned}$$

Robust PCA

- Fault detection rate (FDR)

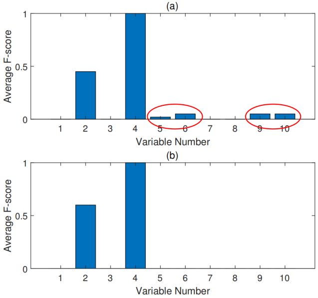
Fault No.	PCA		RPCA		SJSPCA		LRPCA	
	T^2	SPE	T^2	SPE	T^2	SPE	T^2	SPE
1	99.13	99.88	99.25	99.88	99.25	99.88	99.25	100
2	98.38	95.75	98.38	98.00	98.38	99.00	98.38	99.75
3	0.88	2.63	1.88	3.25	2.25	4.25	3.88	6.75
4	20.88	100	30.25	100	34.50	100	39.38	100
5	24.13	20.88	28.75	24.25	30.25	24.25	34.50	24.50
6	99.13	100	99.25	100	99.38	100	99.38	100
7	100	100	100	100	100	100	100	100
8	96.88	83.63	97.13	91.50	97.13	96.75	98.50	96.75
9	1.75	1.75	2.63	2.50	1.75	2.50	3.75	2.75
10	29.63	25.75	33.13	32.75	34.00	30.38	36.13	34.58
11	40.63	74.88	46.37	81.25	48.38	84.88	49.50	90.25
12	98.38	89.50	98.50	90.75	98.50	90.75	99.25	90.75
13	93.63	95.25	93.63	96.25	93.63	97.50	93.63	99.75
14	99.25	100	99.50	100	99.88	100	99.88	100
15	1.38	3.00	2.50	3.88	1.50	3.88	3.75	7.25
16	13.50	27.38	14.13	32.25	14.63	39.50	16.78	39.50
17	76.25	95.38	78.00	95.88	83.50	96.25	88.63	96.75
18	89.25	90.13	89.38	91.25	89.38	92.50	90.63	92.50
19	14.13	18.50	16.25	24.38	16.25	22.38	16.75	28.47
20	31.75	49.75	42.13	52.25	39.38	68.25	48.38	69.63
21	39.25	47.25	39.50	47.38	44.63	49.25	45.75	52.38
Average	55.63	60.14	57.64	62.82	58.41	64.98	**60.29**	**66.75**

- Monitoring results for Fault 10



Sparse constrained PCA

- Xiu-Yang-Kong-Liu, DDCLS, 2020 / Xiu-Miao-Liu, IEEE TII, 2023

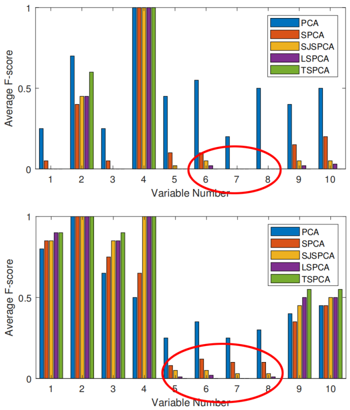


$$\min_{A,B} \quad \frac{1}{2}\|X - XBA^\top\|_F^2 + \lambda_1 \|B\|_{2,1}$$
$$\text{s.t.} \quad A^\top A = I$$
$$\Downarrow$$
$$\min_{A,B} \quad \frac{1}{2}\|X - XBA^\top\|_F^2 + \lambda \mathrm{Tr}(B^\top L B)$$
$$\text{s.t.} \quad A^\top A = I, \ \|B\|_{2,0} \le s$$

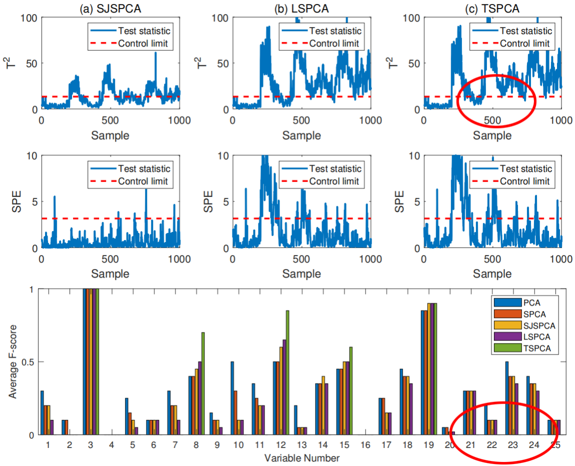
- Alternating direction method of multipliers (ADMM)
- Two-stage monitoring framework
  - Perform fault detection using residual generators
  - Do fault isolation by shrinking the sparsity level s

Sparse constrained PCA

- Simulation examples



- Application on the cylinder-piston process



Sparse CCA

- Xiu-Yang-Kong-Liu, TCSII, 2021

$$\begin{aligned} \min_{A,B} \quad & \frac{1}{2}\|XA - YB\|_F^2 \\ \text{s.t.} \quad & A^\top X^\top XA = I,\ B^\top Y^\top YB = I \end{aligned}$$

$$\Downarrow$$

$$\begin{aligned} \min_{A,B} \quad & \frac{1}{2}\|XA - YB\|_F^2 + \lambda_1\|A\|_{2,1} + \lambda_2\|B\|_{2,1} \\ \text{s.t.} \quad & A^\top X^\top XA = I,\ B^\top Y^\top YB = I \end{aligned}$$

$$\Downarrow$$

$$\begin{aligned} \min_{A,B} \quad & \frac{1}{2}\|XA - YB\|_F^2 + \lambda_1\|A\|_{2,1} + \lambda_2\|B\|_{2,1} + \mu_1 \mathrm{Tr}(A^\top L_1 A) + \mu_2 \mathrm{Tr}(B^\top L_2 B) \\ \text{s.t.} \quad & A^\top X^\top XA = I,\ B^\top Y^\top YB = I \end{aligned}$$

- Alternating minimization algorithm (AMA)
- The generated sequence $\{(A^k, B^k)\}$ converges to a local minimizer

Sparse CCA

- Offline modeling
  - Normalize the training datasets
  - Compute the projections using SISCCA
  - Determine the control limit and construct detection logic
- Online monitoring
  - Normalize the testing datasets
  - Calculate the monitoring statistics
  - Make a decision according to the detection logic

- Monitoring results of FDR and FAR

Fault No.	CCA		SCCA		JSCCA		SJSCCA	
	FDR	FAR	FDR	FAR	FDR	FAR	FDR	FAR
1	99.25	**0.00**	99.38	**0.00**	99.50	**0.00**	**99.75**	**0.00**
2	98.62	0.63	**99.47**	**0.00**	**99.47**	**0.00**	**99.47**	**0.00**
3	33.80	2.50	36.64	1.88	38.20	**0.00**	**41.38**	**0.00**
4	**100**	1.88	**100**	0.63	**100**	**0.00**	**100**	**0.00**
5	29.63	1.88	31.00	0.63	34.50	**0.00**	**36.62**	**0.00**
6	99.88	0.63	99.90	**0.00**	**100**	**0.00**	**100**	**0.00**
7	**100**	1.88	**100**	0.63	**100**	0.63	**100**	**0.00**
8	93.25	1.88	95.00	0.63	95.26	**0.00**	**97.85**	**0.00**
9	31.20	3.13	35.25	2.50	38.50	**0.63**	**40.87**	**0.63**
10	27.50	1.25	32.62	**0.00**	36.13	**0.00**	**39.58**	**0.00**
11	66.37	0.63	69.91	**0.00**	72.00	**0.00**	**78.50**	**0.00**
12	90.75	1.25	93.87	0.63	94.50	0.63	**96.37**	**0.00**
13	91.37	0.63	92.00	0.63	93.67	**0.00**	**95.29**	**0.00**
14	85.00	1.88	86.50	**0.63**	88.12	**0.63**	**89.82**	**0.63**
15	36.20	3.13	39.57	1.25	40.84	0.63	**42.37**	**0.00**
16	15.88	7.50	19.13	4.38	22.75	3.13	**26.37**	**1.25**
17	33.37	3.13	36.00	3.13	37.25	3.13	**41.75**	**2.50**
18	87.88	1.88	89.70	0.63	91.56	0.63	**94.12**	**0.00**
19	22.25	**1.25**	25.66	**1.25**	27.08	**1.25**	**29.93**	**1.25**
20	49.63	0.63	51.80	**0.00**	**55.75**	**0.00**	**55.75**	**0.00**
21	90.00	1.25	91.75	1.25	93.63	**0.63**	**96.60**	**0.63**
Average	65.80	1.85	67.86	0.98	69.46	0.57	**71.54**	**0.33**

Sparse constrained CCA

- Xiu-Miao-Liu, IEEE TNNLS, 2024

$$
\begin{aligned}
\min_{A,B}\quad & -\operatorname{Tr}(A^\top X^\top YB)\\
\text{s.t.}\quad & A^\top X^\top XA = I,\ B^\top Y^\top YB = I
\end{aligned}
$$

$$\Downarrow$$

$$
\begin{aligned}
\min_{A,B}\quad & -\operatorname{Tr}(A^\top X^\top YB)\\
\text{s.t.}\quad & A^\top X^\top XA = I,\ B^\top Y^\top YB = I\\
& \|A\|_{2,0} \le s_1,\ \|B\|_{2,0} \le s_2
\end{aligned}
$$

- Alternating minimization algorithm (AMA) + Manifold optimization

$$
\begin{aligned}
\min_{A,B,C,D}\quad & -\frac{1}{N}\operatorname{Tr}(C^\top D) + \frac{\beta}{2}\|XA - C\|_F^2 + \frac{\beta}{2}\|YB - D\|_F^2\\
\text{s.t.}\quad & \|A\|_{2,0} \le s_1,\ \|B\|_{2,0} \le s_2\\
& C^\top C = I,\ D^\top D = I
\end{aligned}
$$

Sparse constrained CCA

- Suppose that $\{A^k\}$ is a generated sequence and X has an upper restricted isometry constant C_{2s_1}. Whenever $0 < \alpha_k \leq \frac{1}{C_{2s_1}+\sigma}$, it holds that

$$G(A^{k+1}) \leq G(A^k) - \frac{\sigma}{2}\|A^{k+1} - A^k\|_F^2.$$

  When $k \to \infty$, it derives that $\|A^{k+1} - A^k\|_F \to 0$ and $\|(\nabla G(A^k))_{supp(A^k)}\|_F \to 0$.
- Suppose that $\{C^k\}$ is a generated sequence. Then there exist $\bar{\gamma}_1 > 0$ and $\bar{\beta} > 0$ such that

$$H(C^{k+1}) - H(C^k) \leq -\bar{\beta}\|V^k\|_F^2.$$

- Suppose that $\{(A^k, B^k, C^k, D^k)\}$ is a sequence generated. Moreover, X and Y satisfy SRIP with constants C_{2s_1}, c_{2s_1} and C_{2s_2}, c_{2s_2}, respectively. Then the sequence converges to a stationary point. Further, our algorithm returns an ϵ-stationary point in at most

$$\lfloor (F(A^0, B^0, C^0, D^0) - F^*)/((\bar{\sigma}_1 + \bar{\sigma}_2 + 2\bar{\beta})\epsilon) \rfloor + 1$$

  iterations, where F^* denotes a lower bound with $\bar{\sigma}_1$, $\bar{\sigma}_2$, and $\bar{\beta}$ being constants.

Kernel CCA

- Xiu-Li, IEEE JSEN, 2023

$$\begin{aligned}\min_{A,B} \quad & -\mathrm{Tr}(A^\top X^\top YB) \\ \text{s.t.} \quad & A^\top X^\top XA = I,\ B^\top Y^\top YB = I \\ & \|A\|_{2,0} \le s_1,\ \|B\|_{2,0} \le s_2\end{aligned}$$

$$\Downarrow$$

$$\begin{aligned}\min_{A,B} \quad & -\mathrm{Tr}(A^\top K_X^\top K_Y B) + \lambda_1 \mathrm{Tr}(A^\top L_1 A) + \lambda_2 \mathrm{Tr}(B^\top L_2 B) \\ \text{s.t.} \quad & A^\top K_X^\top K_X A = I,\ B^\top K_Y^\top K_Y B = I \\ & \|A\|_{2,0} \le s_1,\ \|B\|_{2,0} \le s_2\end{aligned}$$

- Note that $K_X = \langle \Phi_X, \Phi_X \rangle = [\kappa_X(\mathbf{x}_i, \mathbf{x}_j)]_{i,j=1}^n$, $K_Y = \langle \Phi_Y, \Phi_Y \rangle = [\kappa_Y(\mathbf{y}_i, \mathbf{y}_j)]_{i,j=1}^n$
- Alternating direction method of multipliers (ADMM)?
- Alternating minimization algorithm (AMA)?

Kernel CCA

- First, compute the classical KCCA to get a good initial point

$$A = \arg\min_{A} \ \frac{1}{2}\|K_X A - T_X\|_F^2$$
$$B = \arg\min_{B} \ \frac{1}{2}\|K_Y B - T_Y\|_F^2$$

- Next, solve the sparse and graph constrained problems
  - Update A

$$\min_{A} \ \frac{1}{2}\|K_X A - T_X\|_F^2 + \lambda_1 \mathrm{Tr}(A^\top L_1 A)$$
$$\text{s.t.} \ \|A\|_{2,0} \leq s_1$$

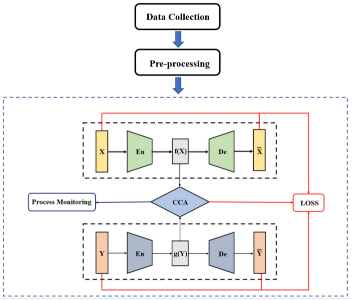
  - Update B

$$\min_{B} \ \frac{1}{2}\|K_Y B - T_Y\|_F^2 + \lambda_2 \mathrm{Tr}(B^\top L_2 B)$$
$$\text{s.t.} \ \|B\|_{2,0} \leq s_2$$

Outline

Deep CCA

- Xiu-Miao-Yang-Liu, IEEE TII, 2022



$$\begin{aligned}\min_{A,B} \quad & -\operatorname{Tr}(A^\top X^\top Y B) \\ \text{s.t.} \quad & A^\top X^\top X A = I,\ B^\top Y^\top Y B = I \\ & \|A\|_{2,0} \le s_1,\ \|B\|_{2,0} \le s_2\end{aligned}$$

$$\Downarrow$$

$$\begin{aligned}\min_{A,B} \quad & -\operatorname{Tr}(A^\top f(X) g(Y)^\top B) \\ \text{s.t.} \quad & A^\top f(X) f(X)^\top A = I,\ B^\top g(Y) g(Y)^\top B = I \\ & \|A\|_{2,0} \le s_1,\ \|B\|_{2,0} \le s_2\end{aligned}$$

- Loss is defined as

$$-\operatorname{Tr}(A^\top f(X) g(Y)^\top B) + \frac{1}{2}\sum_{i=1}^{N} \|\mathbf{x}_i - p(f(\mathbf{x}_i))\|^2 + \frac{1}{2}\sum_{i=1}^{N} \|\mathbf{y}_i - q(g(\mathbf{y}_i))\|^2$$

Deep CCA

Monitoring results of FDR and FAR

Fault No.	CCA		CCA-SCO		KCCA		KCCA-SCO		DCCA		DCCA-SCO	
	FDR	FAR	FDR	FAR	FDR	FAR	FDR	FAR	FDR	FAR	FDR	FAR
IDV(1)	99.75%	0.63%	99.75%	0.63%	99.88%	0.63%	99.88%	0.00%	99.88%	0.00%	99.88%	0.00%
IDV(2)	96.50%	0.63%	97.25%	0.63%	98.38%	0.00%	98.38%	0.00%	98.47%	0.00%	99.50%	0.00%
IDV(4)	100%	1.88%	100%	1.25%	100%	1.25%	100%	0.63%	100%	0.63%	100%	0.00%
IDV(5)	100%	3.75%	100%	3.25%	100%	2.50%	100%	2.50%	100%	2.50%	100%	1.88%
IDV(6)	100%	4.38%	100%	4.38%	100%	3.75%	100%	3.25%	100%	3.25%	100%	3.25%
IDV(7)	100%	3.75%	100%	3.25%	100%	2.50%	100%	1.75%	100%	1.25%	100%	0.63%
IDV(8)	96.50%	1.88%	97.50%	1.88%	97.85%	0.63%	98.25%	0.63%	98.88%	0.63%	99.38%	0.00%
IDV(10)	86.88%	1.25%	87.75%	0.63%	89.58%	0.63%	89.75%	0.63%	90.38%	0.00%	93.88%	0.00%
IDV(11)	76.50%	0.63%	77.50%	0.63%	78.50%	0.63%	79.63%	0.63%	80.13%	0.63%	84.50%	0.00%
IDV(12)	99.00%	1.25%	99.00%	0.63%	99.25%	0.00%	99.37%	0.00%	99.50%	0.63%	99.75%	0.00%
IDV(13)	95.75%	0.63%	96.13%	0.63%	96.50%	0.63%	96.50%	0.63%	96.75%	0.00%	96.88%	0.00%
IDV(14)	100%	1.88%	100%	1.25%	100%	0.63%	100%	0.63%	100%	0.63%	100%	0.63%
IDV(16)	93.00%	7.50%	94.38%	5.63%	95.63%	1.25%	96.63%	1.25%	96.63%	1.25%	98.75%	0.63%
IDV(17)	94.13%	3.13%	94.13%	2.50%	94.25%	2.50%	95.13%	1.75%	95.75%	1.25%	96.38%	1.25%
IDV(18)	90.88%	1.88%	91.25%	1.25%	92.50%	0.00%	92.75%	0.00%	93.50%	0.63%	95.63%	0.00%
IDV(19)	92.00%	1.25%	92.63%	1.25%	94.25%	1.25%	94.25%	0.63%	94.93%	0.63%	95.50%	0.63%
IDV(20)	86.88%	0.63%	87.13%	0.63%	87.75%	0.63%	87.75%	0.00%	88.88%	1.25%	89.38%	0.00%

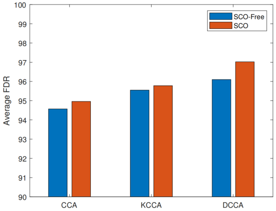
Detection time

CCA	CCA-SCO	KCCA	KCCA-SCO	DCCA	DCCA-SCO
0.044	0.072	15.676	22.556	2.806	3.097

Monitoring performance

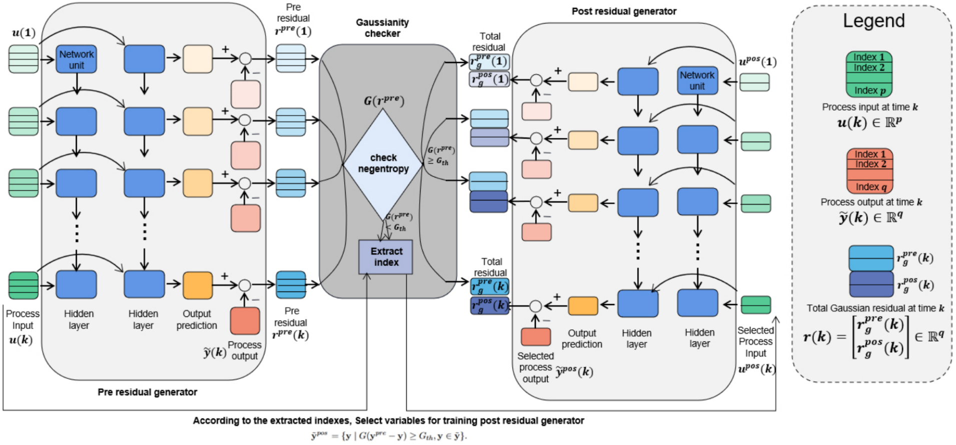
Hidden layers	1	2	3
FDR	89.75%	93.88%	93.88%
FAR	0.63%	0.00%	0.00%

Average comparison



Dual RNN

- Xiu-Zhang-Guo-Liu-Yang, IEEE TIM, 2024

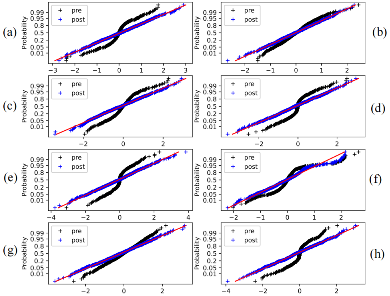


Dual RNN

- Monitoring results of FDR

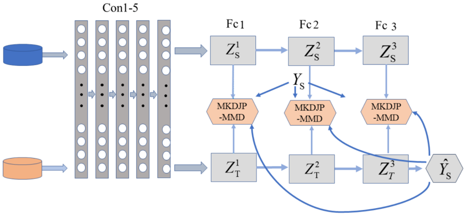
No.	DPCA	DKPCA	SAM	LQ-SAM	RIKPCA	KLD	RNN	D-RNN	D-LSTM	D-GRU
1	99.13%	99.25%	99.50%	99.50%	99.13%	99.25%	99.50%	**99.75%**	99.50%	**99.75%**
2	98.50%	98.50%	97.38%	97.50%	98.50%	98.50%	98.25%	98.50%	**98.63%**	98.50%
3	0.38%	0.38%	2.38%	2.50%	3.63%	3.75%	2.50%	**4.63%**	3.25%	2.25%
4	0.88%	1.75%	24.00%	25.25%	2.75%	22.50%	99.63%	99.75%	**100.00%**	99.63%
5	24.38%	23.63%	28.88%	31.75%	21.00%	27.00%	**32.50%**	26.63%	27.88%	31.50%
6	99.75%	98.88%	**100.00%**	**100.00%**	**100.00%**	**100.00%**	**100.00%**	**100.00%**	**100.00%**	**100.00%**
7	35.50%	36.50%	99.13%	99.50%	38.75%	98.75%	**100.00%**	**100.00%**	**100.00%**	**100.00%**
8	97.25%	97.13%	**97.63%**	97.25%	97.50%	97.50%	95.63%	**97.63%**	96.75%	96.88%
9	0.38%	1.13%	0.88%	1.25%	3.86%	4.00%	1.38%	**4.50%**	2.25%	1.75%
10	25.13%	26.63%	43.00%	43.13%	29.38%	44.50%	27.38%	**60.75%**	43.13%	49.75%
11	10.86%	5.50%	45.00%	49.63%	6.50%	33.63%	61.13%	52.63%	**63.00%**	60.75%
12	99.00%	99.00%	**99.63%**	99.00%	99.00%	**99.63%**	88.38%	99.00%	93.88%	93.00%
13	94.75%	94.25%	99.63%	**99.75%**	95.50%	95.63%	93.25%	95.13%	93.63%	93.25%
14	99.75%	93.13%	99.38%	99.63%	94.75%	99.75%	89.00%	**100.00%**	87.88%	92.63%
15	0.50%	0.75%	2.00%	3.50%	**4.25%**	4.13%	1.63%	4.13%	2.63%	3.50%
16	13.86%	14.25%	20.63%	21.50%	14.25%	21.50%	17.75%	**30.88%**	20.50%	27.13%
17	81.38%	75.75%	79.88%	83.38%	83.25%	90.25%	53.25%	**94.50%**	84.38%	88.13%
18	88.88%	89.25%	95.88%	**96.13%**	90.75%	95.75%	87.63%	90.00%	88.00%	88.25%
19	11.88%	5.88%	20.38%	21.63%	11.75%	25.75%	15.13%	**29.75%**	21.50%	24.25%
20	20.50%	20.13%	43.75%	45.88%	21.50%	44.50%	49.50%	51.50%	51.63%	**52.00%**
21	42.00%	36.25%	39.00%	42.63%	40.75%	**56.75%**	28.25%	49.50%	34.50%	39.25%
Ave.	49.74%	48.52%	58.96%	60.06%	50.31%	58.12%	59.13%	**66.15%**	62.52%	63.91%

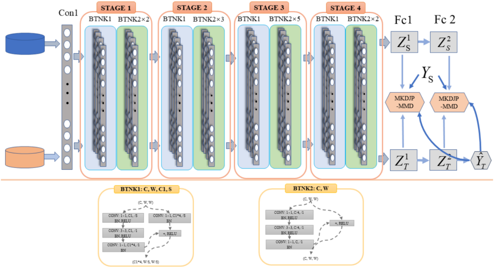
- Probability plots



Deep TL

- Yu, Master's Thesis, 2024



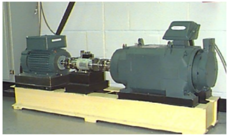


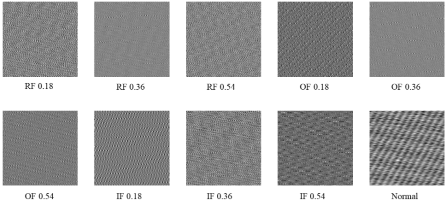
- Deep joint probability adaptation network (DJPAN) + AlexNet / ResNet50
- Loss is defined as

$$\min_{\Theta} \frac{1}{n}\sum_{i=1}^{n} J(\theta(\mathbf{x}_{S,i}), y_{S,i}) + \lambda \sum_{l\in\mathcal{L}} d_{\mathcal{L}}(\mathcal{D}_S^l, \mathcal{D}_T^l)$$

Deep TL

▶ Data processing





▶ Transfer accuracy

Tasks	Methods							
	AlexNet	DDC	DeepCoral	DAN	DANN	DSAN	BNM	DJPAN
SP0 → SP1	74.82	75.18	73.57	82.50	88.84	74.91	88.21	**90.98**
SP0 → SP2	71.09	72.13	77.12	78.06	81.45	82.11	85.78	**89.74**
SP0 → SP3	76.01	76.28	72.99	74.08	81.32	**84.25**	71.25	78.93
SP1 → SP0	44.42	47.38	59.68	78.74	**88.27**	75.69	73.98	81.98
SP1 → SP2	64.41	68.64	68.08	85.40	**99.44**	92.09	82.77	96.23
SP1 → SP3	67.22	77.47	74.48	83.42	86.45	84.98	**87.73**	86.27
SP2 → SP0	58.91	61.20	65.59	67.21	69.30	64.92	69.49	**81.22**
SP2 → SP1	70.89	75.36	78.30	92.41	**96.25**	85.71	91.96	95.45
SP2 → SP3	68.32	75.27	85.35	90.11	81.32	87.82	79.58	**90.29**
SP3 → SP0	61.98	64.92	65.97	71.12	73.21	72.16	69.59	**79.03**
SP3 → SP1	57.86	62.59	62.41	68.30	72.14	61.88	**74.73**	70.89
SP3 → SP1	52.26	54.71	54.33	75.33	**80.13**	67.14	76.46	74.20
Average	64.02	67.59	69.82	78.89	83.18	77.81	79.29	**84.60**

Outline

Future Work

- Deep CCA for FD
  - Chen-Liang-Ding-Yang-Peng-Yuan, A comparative study of deep neural network-aided canonical correlation analysis-based process monitoring and fault detection methods, IEEE TNNLS, 2022
  - Song-Zheng-Jin-Shi-Tao-Tan, A fault-targeted gated recurrent unit-canonical correlation analysis method for incipient fault detection, IEEE TII, 2024
- Deep transfer learning for FD
  - Zhao-Zhang-Yu-Sun-Wang-Yan-Chen, Applications of unsupervised deep transfer learning to intelligent fault diagnosis: A survey and comparative study, IEEE TIM, 2021
  - Chen-Yang-Xue-Huang-Ferrero-Wang, Deep transfer learning for bearing fault diagnosis: A systematic review since 2016, IEEE TIM, 2023
- Large language models for FD
  - Zheng-Pan-Liu-Chen, Empirical study on fine-tuning pre-trained large language models for fault diagnosis of complex systems, RESS, 2024
  - Zhang-Xu-Li-Sun-Bao-Zhang, LLM-TSFD: An industrial time series human-in-the-loop fault diagnosis method based on a large language model, ESA, 2025
  - Tao-Liu-Ning-Cao-Huang-Lu, LLM-based framework for bearing fault diagnosis, MSSP, 2025

References

- Xiu-Pan-Yang-Liu, Efficient and fast joint sparse constrained canonical correlation analysis for fault detection, IEEE Transactions on Neural Networks and Learning Systems, 2024
- Xiu-Zhang-Guo-Liu-Yang, A new end-to-end monitoring framework for nonlinear dynamic processes with unknown noise statistics, IEEE Transactions on Instrumentation and Measurement, 2024
- Xiu-Miao-Liu, A sparsity-aware fault diagnosis framework focusing on accurate isolation, IEEE Transactions on Industrial Informatics, 2023
- Xiu-Li, Learning sparse kernel CCA with graph priors for nonlinear process monitoring, IEEE Sensors Journal, 2023
- Xiu-Miao-Yang-Liu, Deep canonical correlation analysis using sparsity constrained optimization for nonlinear process monitoring, IEEE Transactions on Industrial Informatics, 2022
- Xiu-Yang-Kong-Liu, Data-driven process monitoring using structured joint sparse canonical correlation analysis, IEEE Transactions on Circuits and Systems II: Express Briefs, 2021
- Xiu-Yang-Kong-Liu, Laplacian regularized robust principal component analysis for process monitoring, Journal of Process Control, 2020