

Robust Classifications of WBM Defect Patterns – Multimodal approach

Taiki AI, Sumika ARIMA, Takumi MAEDA, Koichi SUMIYA, Taiki ITO, Daisuke TAKADA. University of Tsukuba

Introduction

- This study proposed a multimodal learning (MML) approach for advanced classifications of defect patterns (DPs) of semiconductor the wafer bin map (WBM) recorded during a final inspection.
- MML approach outperforms previous studies with progressed features or networks only on the WBM-DP data.

	- 01				CLAO	
	Wu's	Nakazawa's	Kyeong's	Wang's	Lee's	Ours
	work	work	work	work	work	Ours
Single type	9	18 *	5	9	5	8=7 +1:None
2 mixed-type		4	6	13	6	21
3 mixed-type		-	4	12	4	35
4 mixed-type		-	1	4	1	35
5 mixed-type		-	-	-	-	21
_						

THE SCOPE OF SUPERPOSITION (SP) CLASSIFICATIONS

Superposed DPs data are generated by data expansion, syntheses and logical operations.



Example of defect pattern of wafer bin map

Causal and random variables are attached to every wafer.

- Causal variables (CVs) (Pc=117): <u>Correct</u> variables that shows causes of defects.
- Random variables (Pr=1100≒10Pc): Incorrect variables that are often included in a real data.

o mixed-type		-	-	-	-	/	
7 mixed-type		-	-	-	-	1	
Total [classes]	9	22	16	38	16	128 ** (=127 + 1:None)	

Methods

Open dataset WM-811k:

- About WM-811k data of WBM DPs of an actual mass production site.
- 1.7 DPs are classified (labeled) \bullet by experts (Right table).

Class Definition											
	Class	# of wafers									
(a)	Center	4294									
(b)	Donut	555									
(c)	Edge-Loc	5189									
(d)	Edge-Ring	9680									
(e)	Loc	3593									
(f)	Near-full	149									
(g)	Random	866									
(h)	Scratch	1193									
	None	147431									
	Total	172950									

Example of causal variables of a pair of the class and cluster

Class	Cluster	Cau	usal variables (Type-A)																		
Class	OldStel	x1	x2	x3	x 4	x5	x6	x7	x8	x9	x10	x11	x12	x13	x14	x15	x16	x17	x18	x19	
Center	3	0	0	0	0	0	0	1	1	1	0	0	0	0	0	0	0	0	0	0	
Center	2	0	0	0	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	•••
Center	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
Scratch	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	
Scratch	2	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	0	0	0	•••
Scratch	1	0	0	0	0	0	0	0	0	0	0	1	1	1	0	0	0	0	0	0	

Example of causal variables 2-mixed SP (e.g. Scratch + Center)

Class	Cluster-1 Cluster-2 Causal variables (Type-A)																					
Class	(Center)	(scratch)	x1	x2	х3	x4	x5	x6	x7	x8	x9	x10	x11	x12	x13	x14	x15	x16	x17	x18	x19	
Scratch + Center	1	3	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	
Scratch + Center	1	2	1	1	1	0	0	0	0	0	0	0	0	0	0	1	1	1	0	0	0	•••
Scratch + Center	1	1	1	1	1	0	0	0	0	0	0	0	1	1	1	0	0	0	0	0	0	
Scratch + Center	2	3	0	0	0	1	1	1	0	0	0	0	0	0	0	0	0	0	1	1	1	
Scratch + Center	2	2	0	0	0	1	1	1	0	0	0	0	0	0	0	1	1	1	0	0	0	•••
Scratch + Center	2	1	0	0	0	1	1	1	0	0	0	0	1	1	1	0	0	0	0	0	0	
Scratch + Center	3	3	0	0	0	0	0	0	1	1	1	0	0	0	0	0	0	0	1	1	1	
Scratch + Center	3	2	0	0	0	0	0	0	1	1	1	0	0	0	1	1	1	0	0	0	0	•••
Scratch + Center	3	1	0	0	0	0	0	0	1	1	1	1	1	1	0	0	0	0	0	0	0	

Results

- All mixed-type and single type (127 classes) are appropriately discriminated by one MML common model.
- MML approach outperforms previous studies (100% column of the table), even with high missing values in CVs and much larger noises.

THE ACCURACY OF PROPOSED METHOD: THE CLASSIFICATION OF ALL SUPERPOSITION DP CLASSES

For multimodal learning (MML) to integrate the defect causal variables, has been proposed and evaluated.

Example of-causal variables added to a pair of a class and a cluster

DP Class	Defect Rate	Cluster No.	x1	x2	x3	x4	x5	x6	x7	x8	x9
Edge-Loc	(0.36, 1.00]	3	1/0	1/0	1/0	1/0	1/0	1/0	1	1	1
Edge-Loc	(0.18, 0.36]	2	1/0	1/0	1/0	1	1	1	0/1	0/1	0/1
Edge-Loc	(0, 0.18]	1	1	1	1	0/1	0/1	0/1	0/1	0/1	0/1

2-step classification and additional MML to achieve classification with higher accuracy for

unbalanced sample data.



- Model of 2-step classification
- Nakazawa's CNN was applied (bottom left Fig) and extended to MML model (bottom right Fig)
- CNN, ViT, and others can be applied.



	# of	Missing rate of causal variables										
	classes	1	0.8	0.6	0.4	0.2	0					
		100%	80%	60%	40%	20%	0%					
Average (1-7 SP all classes)	127	0.521	0.758	0.879	0.952	0.987	1.000					
Average (only single class)	7	0.874	0.964	0.993	0.998	0.999	1.000					
Average (only 2-SP classes)	21	0.656	0.871	0.944	0.980	0.996	1.000					
Average (only 3-SP classes)	35	0.342	0.696	0.882	0.961	0.994	1.000					
Average (only 4-SP classes)	35	0.150	0.535	0.849	0.943	0.994	1.000					
Average (only 5-SP classes)	21	0.214	0.521	0.777	0.940	0.992	1.000					
Average (only 6-SP classes)	7	0.143	0.599	0.817	0.940	0.987	1.000					
Average (6-SP and 7-SP classes)	8	0.225	0.588	0.840	0.951	0.992	1.000					

Conclusion

- The possibilities are clarified to the robust classifications of uncertain numbers of mixedtype DPs by proposed MML approach.
- WM-811K has no superpositions labeled data. However, every device production site has data of the WBM and candidates of DPs' causal variables and can apply the MML, even with much noises.

Sumika Arima: arima@sk.tsukuba.ac.jp Taiki Ai: s2320405@u.tsukuba.ac.jp