Mixed-type Wafer Map Defect Dataset

Description

Defect pattern recognition (DPR) of wafer maps, especially the mixed-type defect, is critical for determining the root cause of production defects.

We collected a large amount of wafer map data in a wafer manufacturing plant. These wafer maps are obtained by testing the electrical performance of each die on the wafer through test probes. However, there are big differences between the quantity distribution of various types of wafer maps actually collected. In order to maintain the balance between the various types of data, we used the confrontation generation network to generate some wafer maps to maintain the balance of the number of samples among the types. Finally, about 38,000 mixed-mode wafer map defect data sets are formed, which are used to identify mixed-mode wafer map defects and assist the research on the causes of defects in the wafer manufacturing process. In order to facilitate researchers, students, and enthusiasts in related fields to better understand the causes of defects in the wafer manufacturing process, we have published a data set of mixed-mode wafer map defects for your reference and research.

Dataset

data_mutil_label.npz

Data

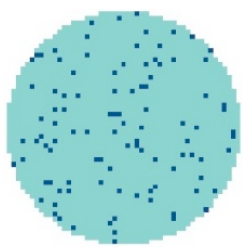
['arr_0']: Defect data of mixed-type wafer map, 0 means blank spot, 1 represents normal die that passed the electrical test, and 2 represents broken die that failed the electrical test. The data shape is 52×52.

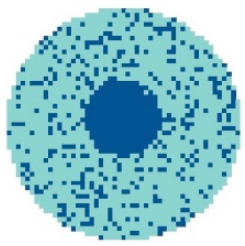
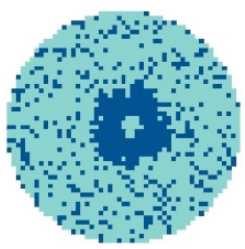
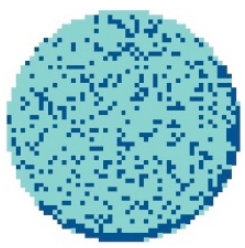
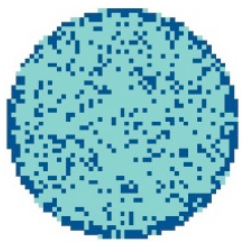
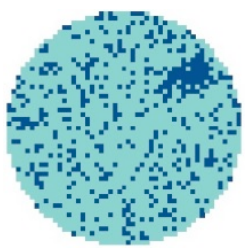
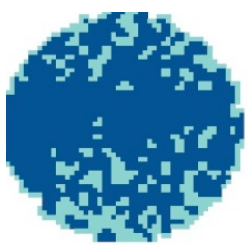
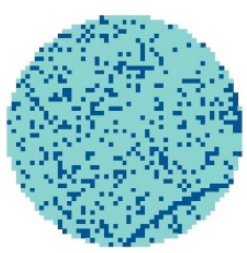
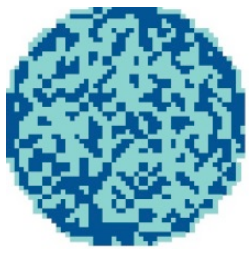
['arr_1']: Mixed-type wafer map defect label, using one-hot encoding, a total of 8 dimensions, corresponding to the 8 basic types of wafer map defects (single defect).

Demonstration

There are 38 types in the mixed-type wafer map defect dataset, including 1 normal type, 8 single defect types, and 29 mixed-type defect types.

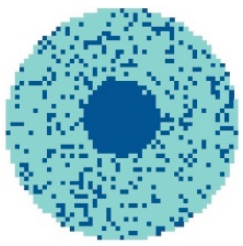
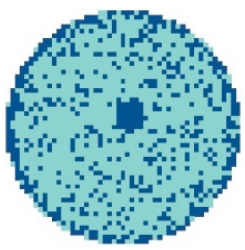
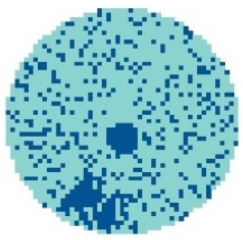
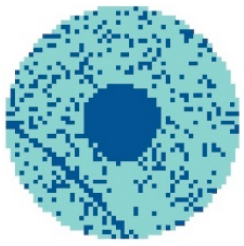
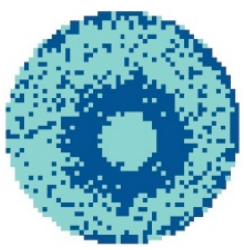
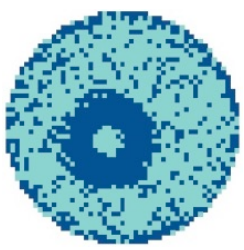
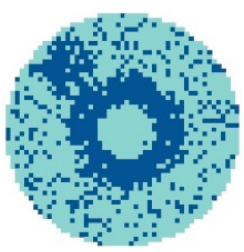
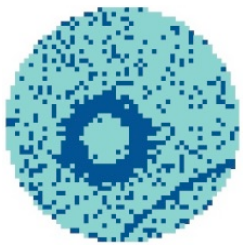
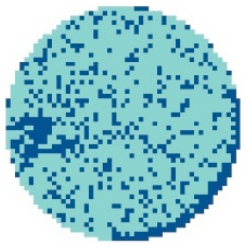
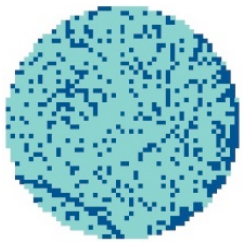
Single Type:

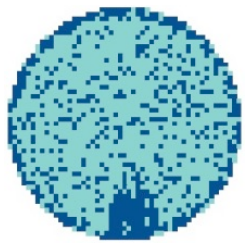
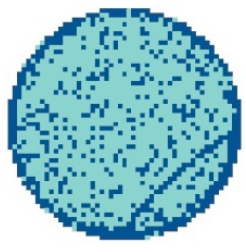
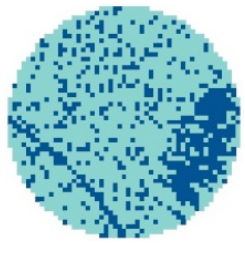
No.	Pattern Type	Pattern Name	Pattern ID	Pattern Label	Illustration	Amount	Index
1	Single-type (9)	Normal	C1	[0 0 0 0 0 0 0 0]		1000	33866-34865

2		Center(C)	C2	[1 0 0 0 0 0 0 0]		1000	12000-12999
3		Donut(D)	C3	[0 1 0 0 0 0 0 0]		1000	24000-24999
4		Edge_Loc(EL)	C4	[0 0 1 0 0 0 0 0]		1000	25000-25999
5		Edge_Ring(ER)	C5	[0 0 0 1 0 0 0 0]		1000	26000-26999
6		Loc(L)	C6	[0 0 0 0 1 0 0 0]		1000	32000-32999
7		Near_Full(NF)	C7	[0 0 0 0 0 1 0 0]		866	33000-33865
8		Scratch(S)	C8	[0 0 0 0 0 0 1 0]		1000	37015-38014
9		Random(R)	C9	[0 0 0 0 0 0 0 1]		149	34866-35014

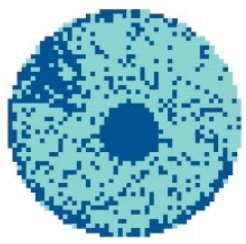
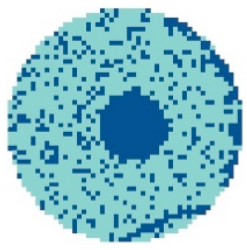
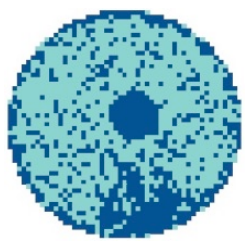
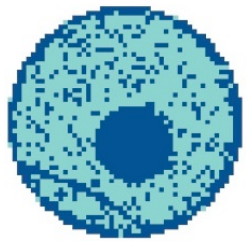
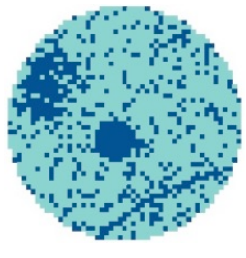
Two Types Mixed:

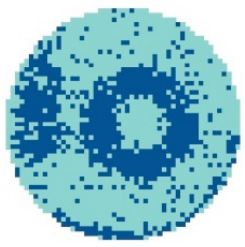
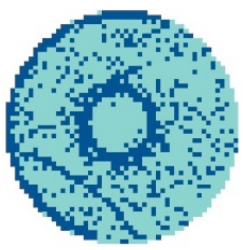
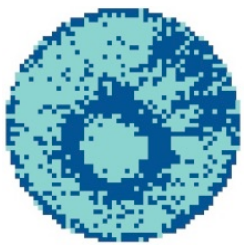
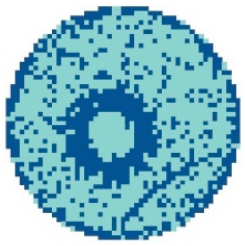
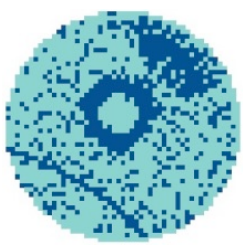
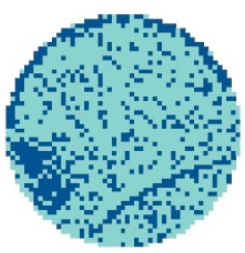
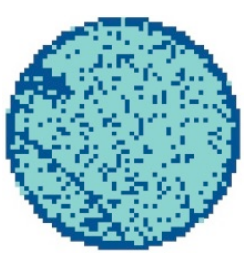
No.	Pattern Type	Pattern Name	Pattern ID	Pattern Label	Illustration	Amount	Index

1	2 mixed-type (13)	C+EL	C10	[1 0 1 0 0 0 0 0]		1000	2000-2999
2		C+ER	C11	[1 0 0 1 0 0 0 0]		1000	4000-4999
3		C+L	C12	[1 0 0 0 1 0 0 0]		1000	10000-10999
4		C+S	C13	[1 0 0 0 0 0 1 0]		1000	11000-11999
5		D+EL	C14	[0 1 1 0 0 0 0 0]		1000	14000-14999
6		D +ER	C15	[0 1 0 1 0 0 0 0]		1000	16000-16999
7		D+L	C16	[0 1 0 0 1 0 0 0]		1000	22000-22999
8		D+S	C17	[0 1 0 0 0 0 1 0]		1000	23000-23999
9		EL+L	C18	[0 0 1 0 1 0 0 0]		1000	28000-28999
10		EL+S	C19	[0 0 1 0 0 0 1 0]		1000	35015-36014

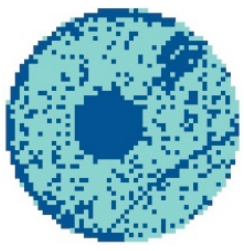
11		ER+L	C20	[0 0 0 1 1 0 0 0]		1000	30000-30999
12		ER+S	C21	[0 0 0 1 0 0 1 0]		1000	36015-37014
13		L+S	C22	[0 0 0 0 1 0 1 0]		1000	31000-31999

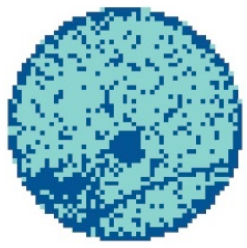
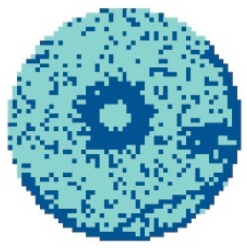
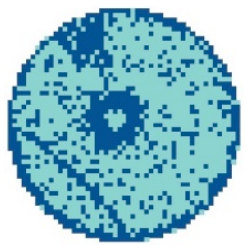
Three Types Mixed:

No.	**Pattern Type**	**Pattern Name**	**Pattern ID**	**Pattern Label**	**Illustration**	**Amount**	**Index**
1	3 mixed-type (12)	C+EL+L	C23	[1 0 1 0 1 0 0 0]		1000	6000-6999
2		C+EL+S	C24	[1 0 1 0 0 0 1 0]		2000	0-1999
3		C+ER+L	C25	[1 0 0 1 1 0 0 0]		1000	8000-8999
4		C+ER+S	C26	[1 0 0 1 0 0 1 0]		1000	3000-3999
5		C+L+S	C27	[1 0 0 0 1 0 1 0]		1000	9000-9999

6		D+EL+L	C28	[0 1 1 0 1 0 0 0]		1000	18000-18999
7		D+EL+S	C29	[0 1 1 0 0 0 1 0]		1000	13000-13999
8		D+ER+L	C30	[0 1 0 1 1 0 0 0]		1000	20000-20999
9		D+ER+S	C31	[0 1 0 1 0 0 1 0]		1000	15000-15999
10		D+L+S	C32	[0 1 0 0 1 0 1 0]		1000	21000-21999
11		EL+L+S	C33	[0 0 1 0 1 0 1 0]		1000	27000-27999
12		ER+L+S	C34	[0 0 0 1 1 0 1 0]		1000	29000-29999

Four Types Mixed:

No.	Pattern Type	Pattern Name	Pattern ID	Pattern Label	Illustration	Amount	Index
1	4 mixed-type (4)	C+L+EL+S	C35	[1 0 1 0 1 0 1 0]		1000	5000-5999

2		C+L+ER+S	C36	[1 0 0 1 1 0 1 0]		1000	7000-7999
3		D+L+EL+S	C37	[0 1 1 0 1 0 1 0]		1000	17000-17999
4		D+L+ER+S	C38	[0 1 0 1 1 0 1 0]		1000	19000-19999

Citation

J. Wang, C. Xu, Z. Yang, J. Zhang and X. Li, "Deformable Convolutional Networks for Efficient Mixed-type Wafer Defect Pattern Recognition," in IEEE Transactions on Semiconductor Manufacturing, doi: 10.1109/TSM.2020.3020985.

Acknowledgement

Thanks to Dr. Uzma Batool from the University of Technology Malaysia for correcting the label errors in the original dataset!

The C7 and C9 labels in the dataset have been corrected!

More

On GitHub:

https://github.com/Junliangwangdhu/WaferMap

On Aliyun:

https://tianchi.aliyun.com/dataset/dataDetail?dataId=77328