Yield Prediction & Enhancement in the Semiconductor Manufacturing

이창현

2018.07.06

목차

- 반도체 제조업 이해
- 반도체 제조 공정 이해
- Defect data 를 이용한 수율 예측 및 원인 공정 탐지
- Defect pattern 을 이용한 Wafer 분류 및 유사 Wafer 검색

반도체 제조업 이해

기업의 목표, 가치

= 이윤의 극대화

반도체 제조업 이해

이윤 = 영업 이익

영업 이익↑ = 매출액↑ – 매출원가↓ – 판매비,관리비↓

- 생산량 증가
- 가격 상승

- 재료비 절감
- 공정수 단축

- 인건비
- 세금, 공과금

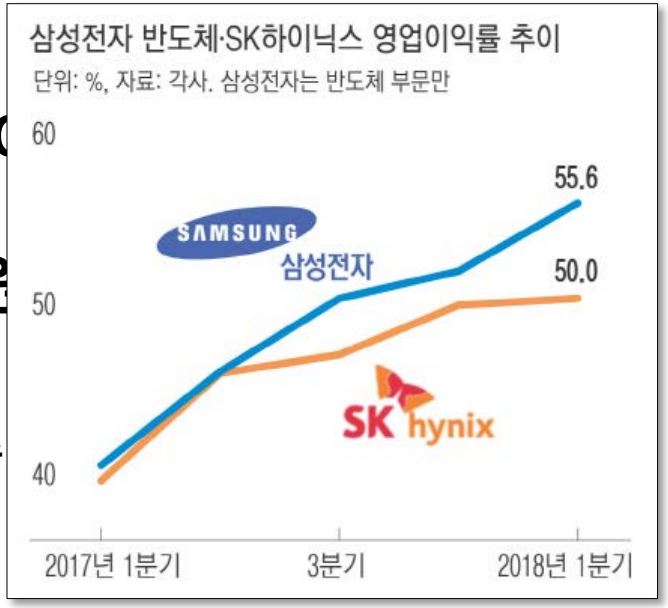
반도체 제조업 이해

이윤 = 영업 (

영업 이익↑ = 매출액↑ – 매출원

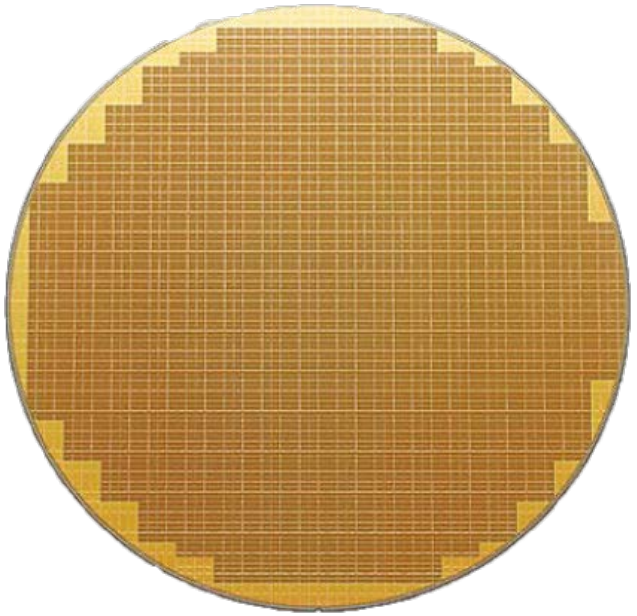
- 생산량 증가
- 가격 상승

- 재료비
- 공정수



반도체 제조업 이해

- 반도체 공정의 기본 단위 = WAFER, LOT



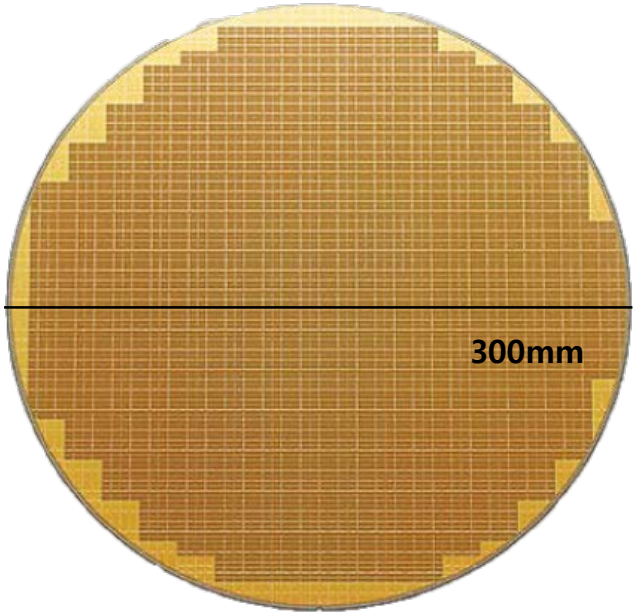
WAFER

1LOT = 25 WFs

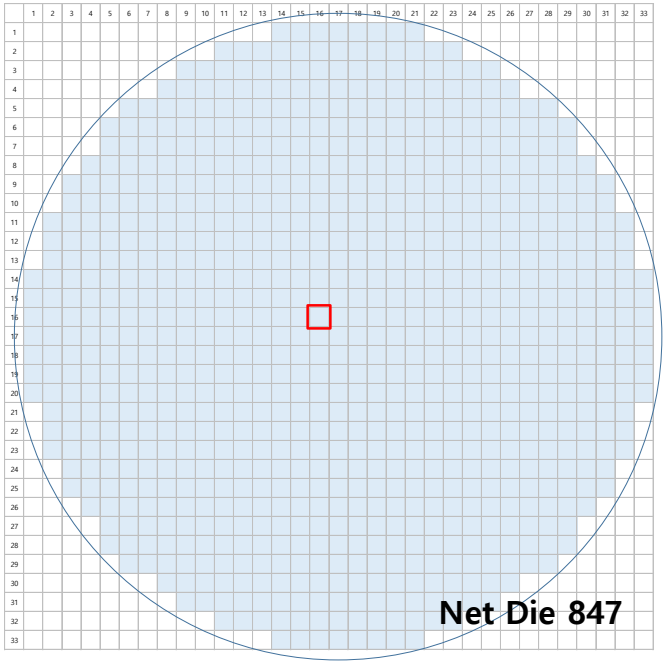
생산 기준 = WAFER 수

반도체 제조업 이해

- **WAFER, DIE(Chip)**



WAFER



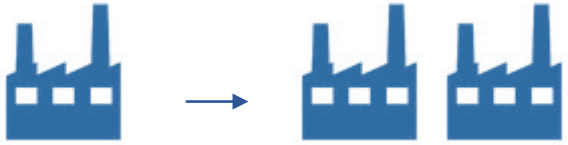
DIE(Chip)

최대한 많은 DIE 를 생산하는 것이 목표

반도체 제조업 이해

- **생산량 증가를 위한 방법**

공장 증설

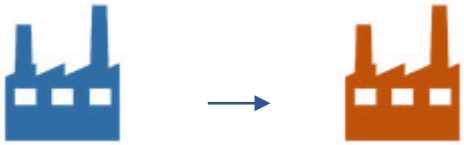


10만장/月 → 20만장/月

공정 불량률 감소

10만장/月 X 0.6 → 10만장/月 X 0.8

공정 시간 단축



10만장/月 → 11만장/月

제품 사이즈 감소

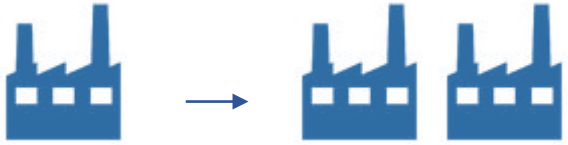
10만장/月 X 1000 → 10만장/月 X 1200

반도체 제조업 이해

- **생산량 증가를 위한 방법**

공장 증설

High Risk

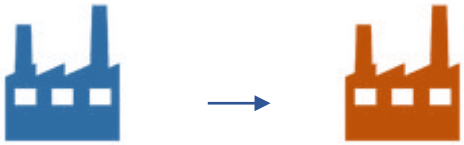


10만장/月 → 20만장/月

공정 불량률 감소

10만장/月 X 0.6 → 10만장/月 X 0.8

공정 시간 단축



10만장/月 → 11만장/月

제품 사이즈 감소

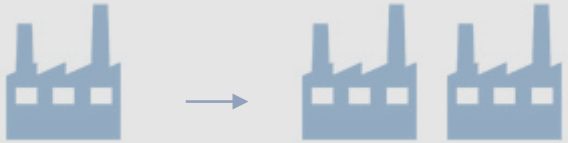
10만장/月 X 1000 → 10만장/月 X 1200

반도체 제조업 이해

- **생산량 증가를 위한 방법**

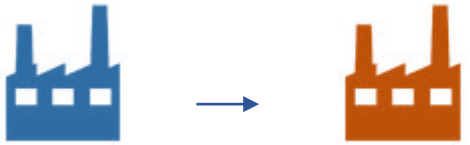
공장 증설

High Risk



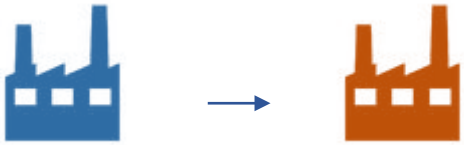
10만장/月 → 20만장/月

공정 불량률 감소



10만장/月 X 0.6 → 10만장/月 X 0.8

공정 시간 단축



10만장/月 → 11만장/月

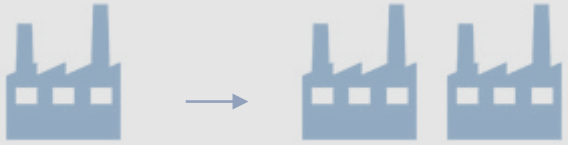
제품 사이즈 감소

10만장/月 X 1000 → 10만장/月 X 1200

반도체 제조업 이해

- **생산량 증가를 위한 방법**

공장 증설 *High Risk*

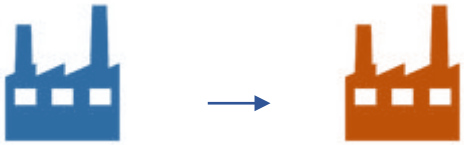


10만장/月 → 20만장/月

공정 불량률 감소

10만장/月 X 0.6 → 10만장/月 X 0.8

공정 시간 단축 *High Tech*



10만장/月 → 11만장/月

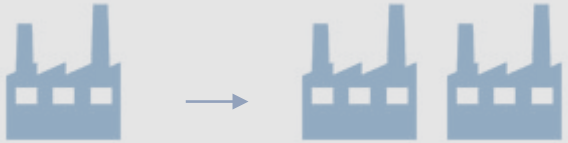
제품 사이즈 감소

10만장/月 X 1000 → 10만장/月 X 1200

반도체 제조업 이해

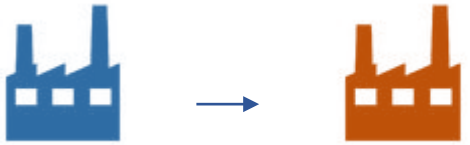
- **생산량 증가를 위한 방법**

공장 증설 *High Risk*



10만장/月 → 20만장/月

공정 불량률 감소



10만장/月 X 0.6 → 10만장/月 X 0.8

공정 시간 단축 *High Tech*

10만장/月 → 11만장/月

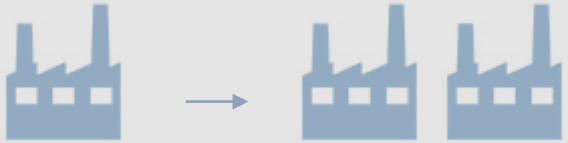
제품 사이즈 감소

10만장/月 X 1000 → 10만장/月 X 1200

반도체 제조업 이해

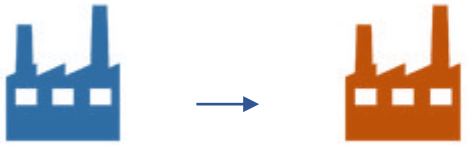
- **생산량 증가를 위한 방법**

공장 증설 *High Risk*



10만장/月 → 20만장/月

공정 불량률 감소



10만장/月 X 0.6 → 10만장/月 X 0.8

공정 시간 단축 *High Tech*

10만장/月 → 11만장/月

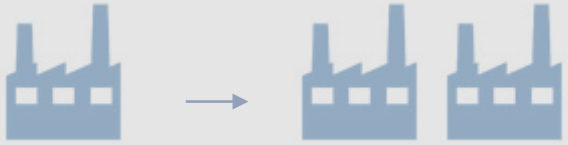
제품 사이즈 감소 *High Tech*

10만장/月 X 1000 → 10만장/月 X 1200

반도체 제조업 이해

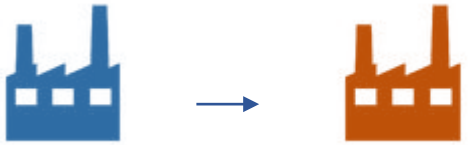
- **생산량 증가를 위한 방법**

공장 증설



10만장/月 → 20만장/月

공정 불량률 감소



10만장/月 X 0.6 → 10만장/月 X 0.8

공정 시간 단축

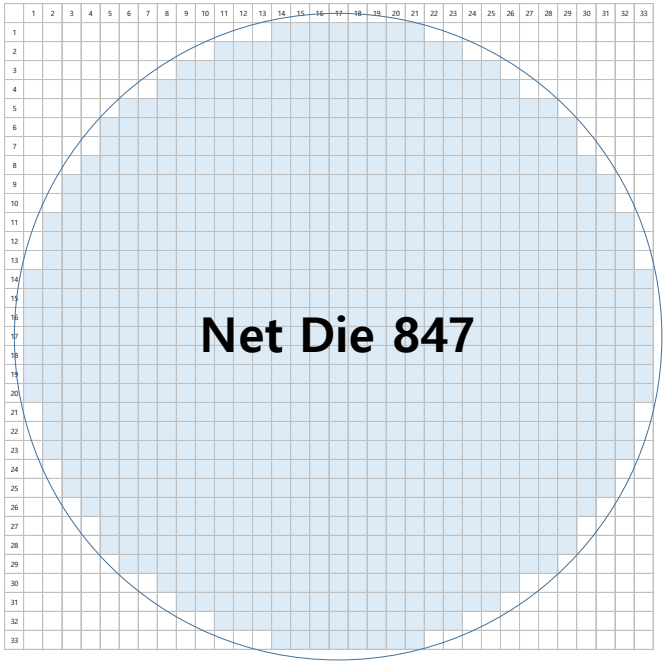
10만장/月 → 11만장/月

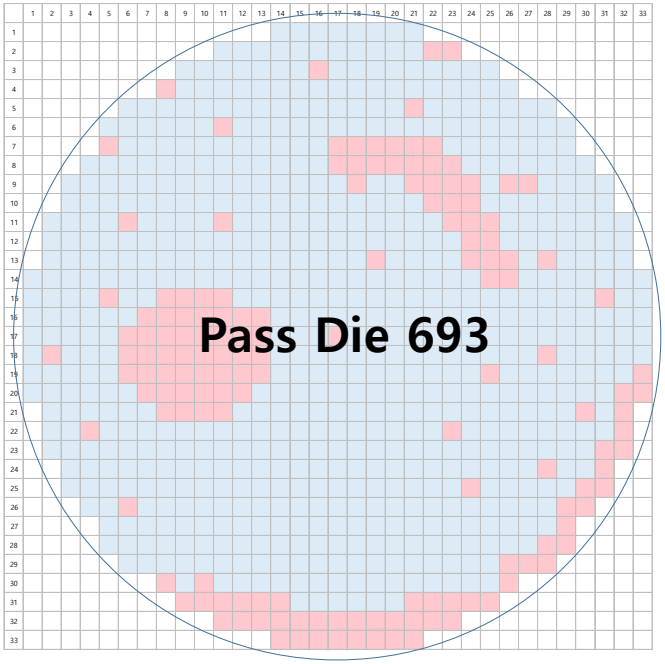
제품 사이즈 감소

10만장/月 X 1000 → 10만장/月 X 1200

반도체 제조업 이해

- **Yield = Pass Die / Net Die**

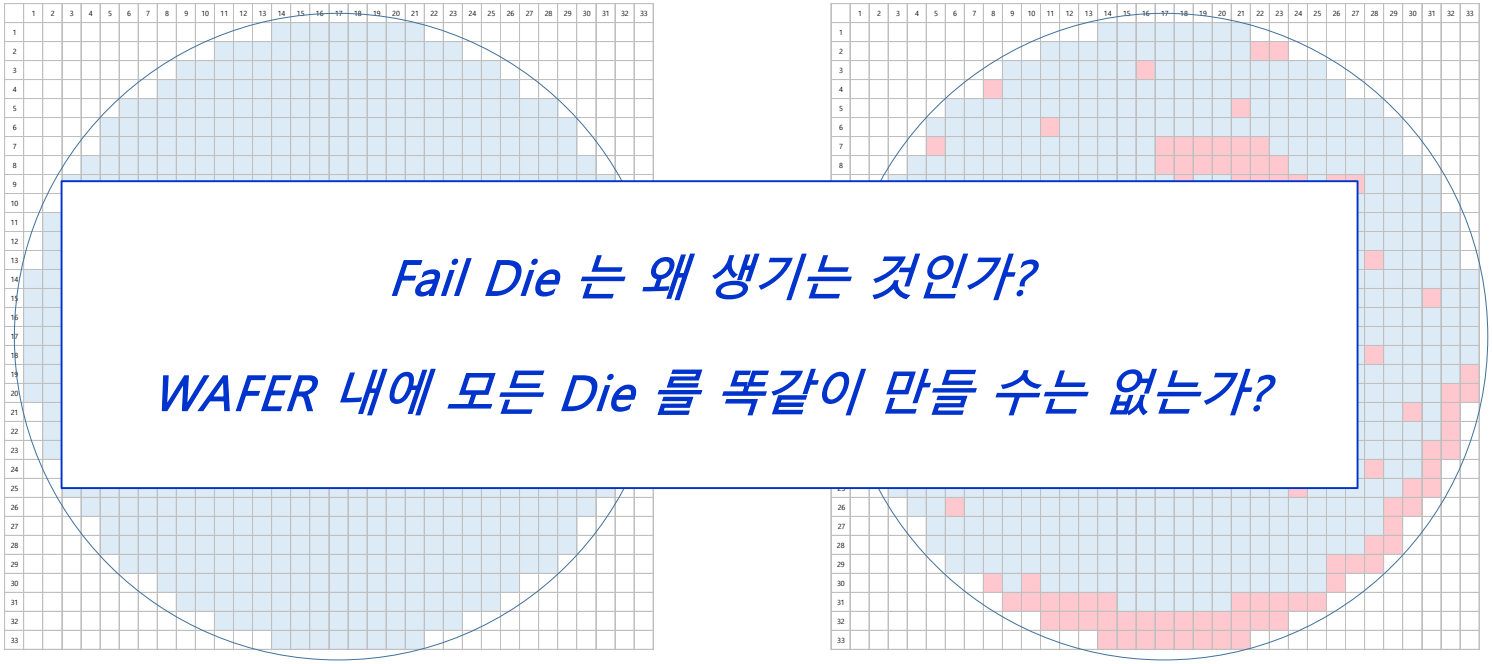




Yield = 693 / 847 = 81.8%

반도체 제조업 이해

- **Yield = Pass Die / Net Die**



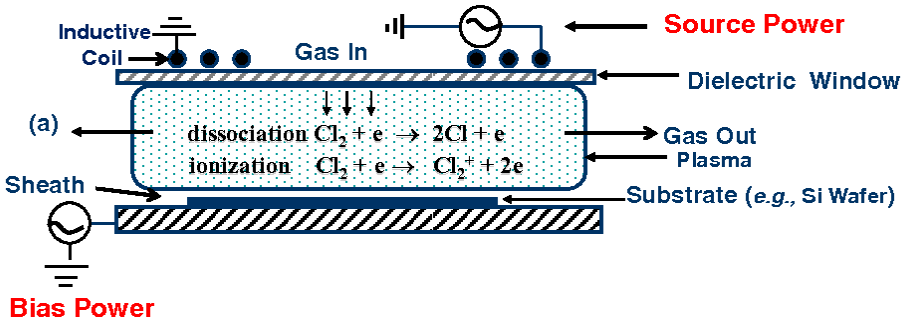
Fail Die 는 왜 생기는 것인가?

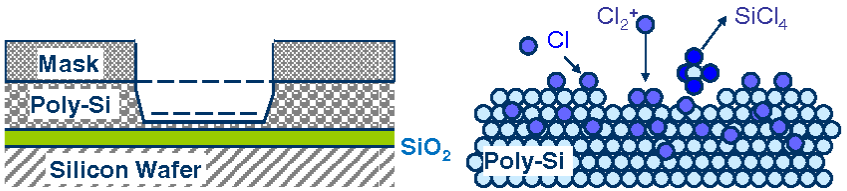
WAFER 내에 모든 Die 를 똑같이 만들 수는 없는가?

Yield = 693 / 847 = 81.8%

반도체 제조 공정 이해

- **ETCH Process 개요**

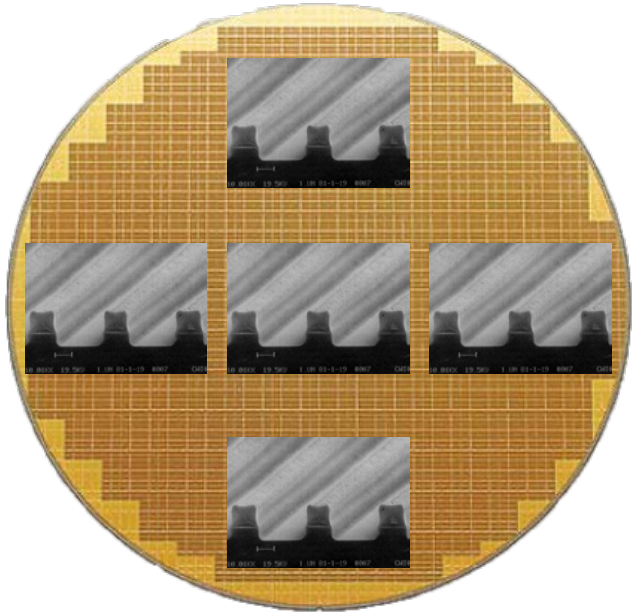




Uniformity & By-product

반도체 제조 공정 이해

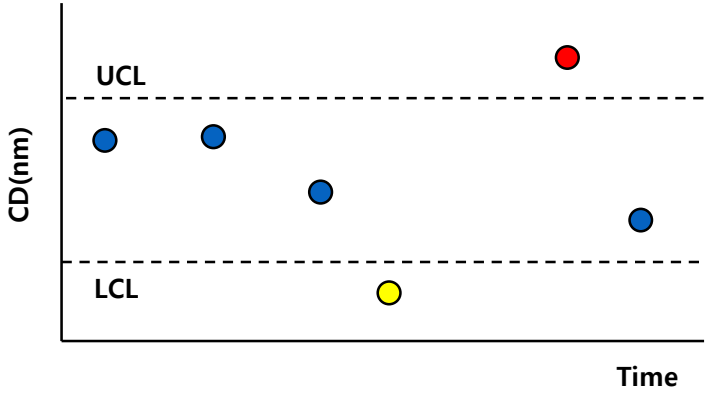
- **ETCH Process – 1. 균일하게 만들기는 힘들다.**



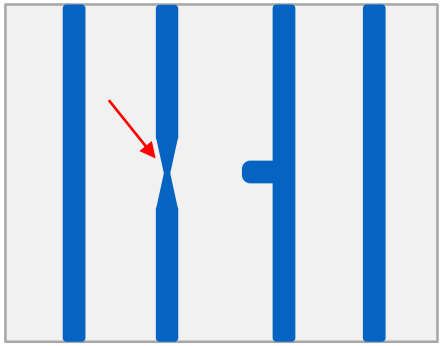
- **WAFER 내 지역적으로 ETCH 차이 발생.**
- **압력, 온도, Gas 농도 등 WAFER 전 영역에서 주변 환경이 미세하게 차이 날 수 밖에는 없음.**
- **이런 미세한 차이로 인하여 지역간 형성되는 Pattern 차이 발생.**
- **이런 차이로 인하여 지역적 불량 발생.**

반도체 제조 공정 이해

- **Metrology – Critical Dimension**

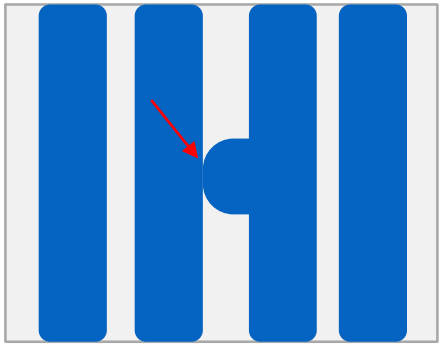


Pattern Metrology 를 이용하여 공정 진행 상태를 점검.
하지만, Wafer 전 영역이나 전체 Wafer 를 점검 하는 것은
불가능함. (시간, 자원 제약) → Virtual Metrology !!!



Open Fail

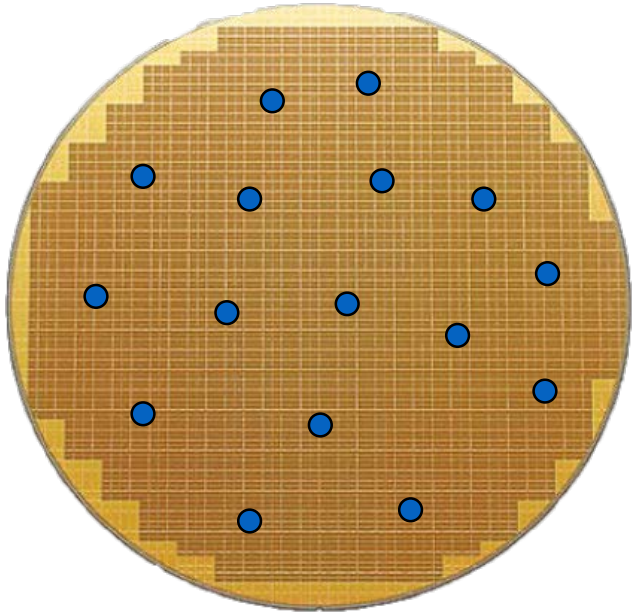
Normal



Short Fail

반도체 제조 공정 이해

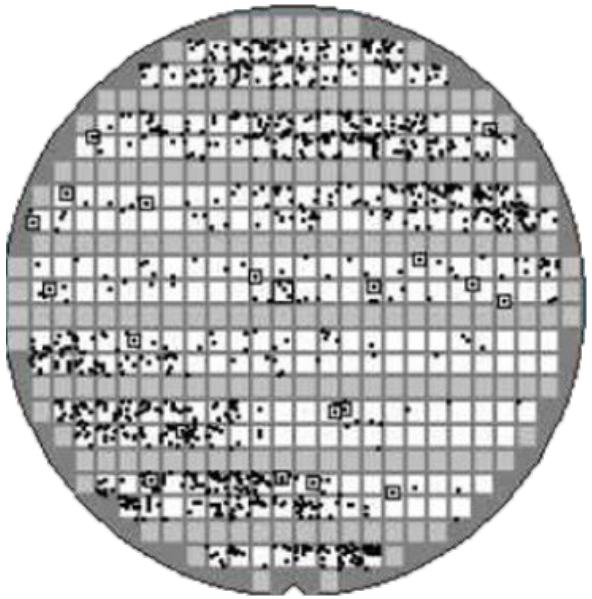
- **ETCH Process – 2. 부산물 처리가 어렵다.**



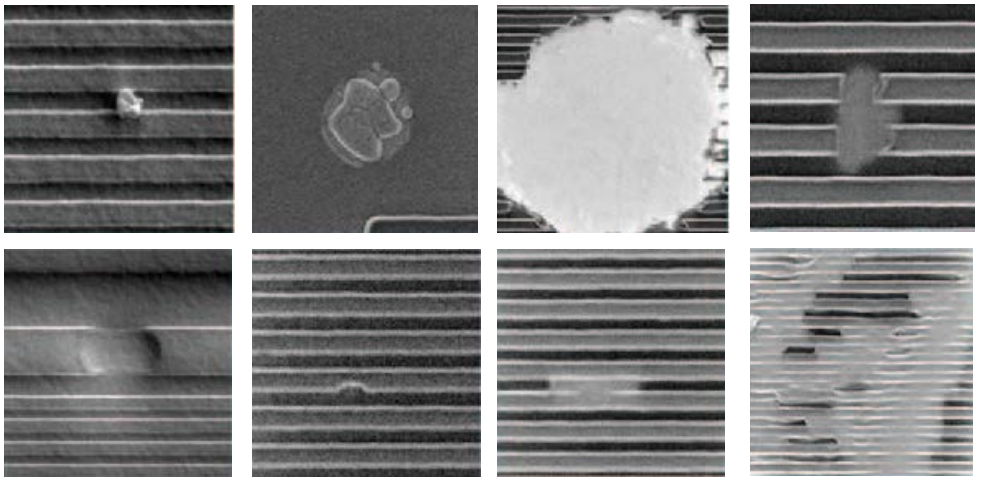
- **ETCH Process 후 발생 된 부산물 제거 필요.**
- **ETCH Process 중 부산물이 잘 빠져나갈 수 있게 환경을 만들어 주는 것이 중요.**
- **잔여 부산물에 대하여 후속 Cleaning Process 에서 한번 더 제거.**
- **완벽히 모든 부산물을 제거 하는 것은 힘듦.**

반도체 제조 공정 이해

- **Inspection – Find Abnormalities**

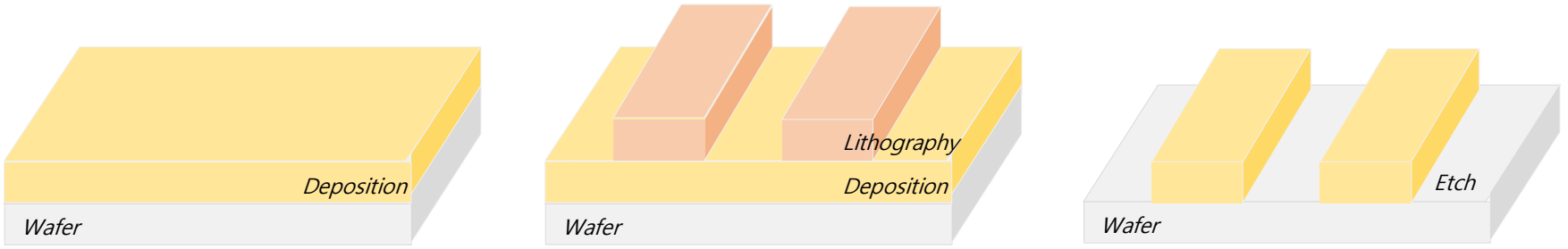


Inspection 장비를 이용하여 Wafer 내 비정상적인 부분 탐색.
탐색 된 IMG 를 바탕으로 해당 공정에 문제가 있는지 확인 가능.
해당 IMG 가 수율과 연관이 있는지 확인 하는 것이 중요.



반도체 제조 공정 이해

- **반도체 기본 공정**

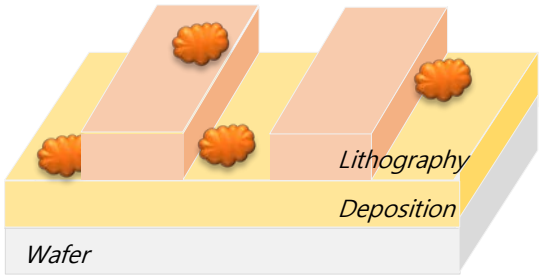


DEP LITHO ETCH

반도체 제조 공정 이해

- 공정 진행 후 부산물 생성

DEP

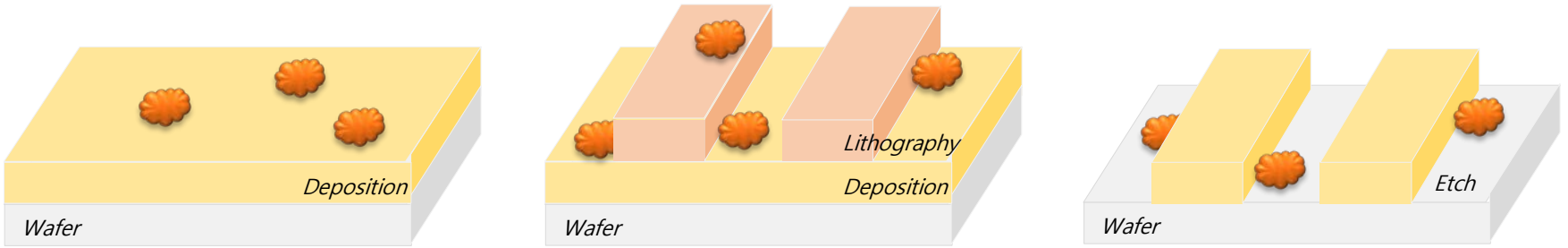


LITHO

ETCH

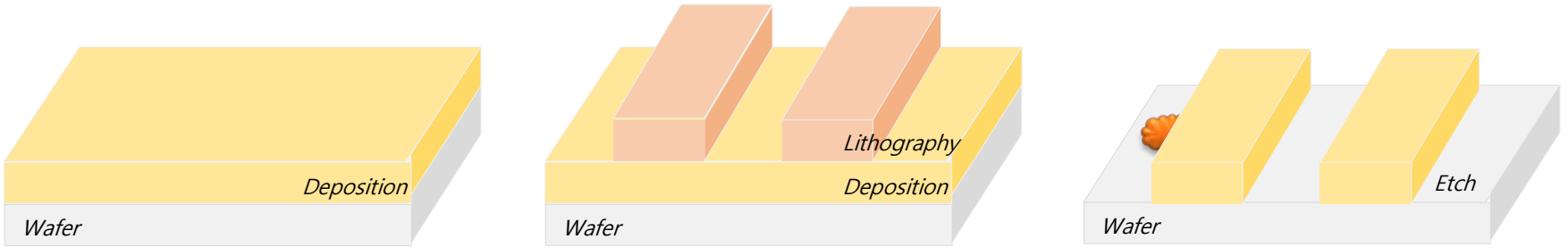
반도체 제조 공정 이해

- 공정 진행 후 생성된 부산물 제거 위하여 CLN 공정 필요



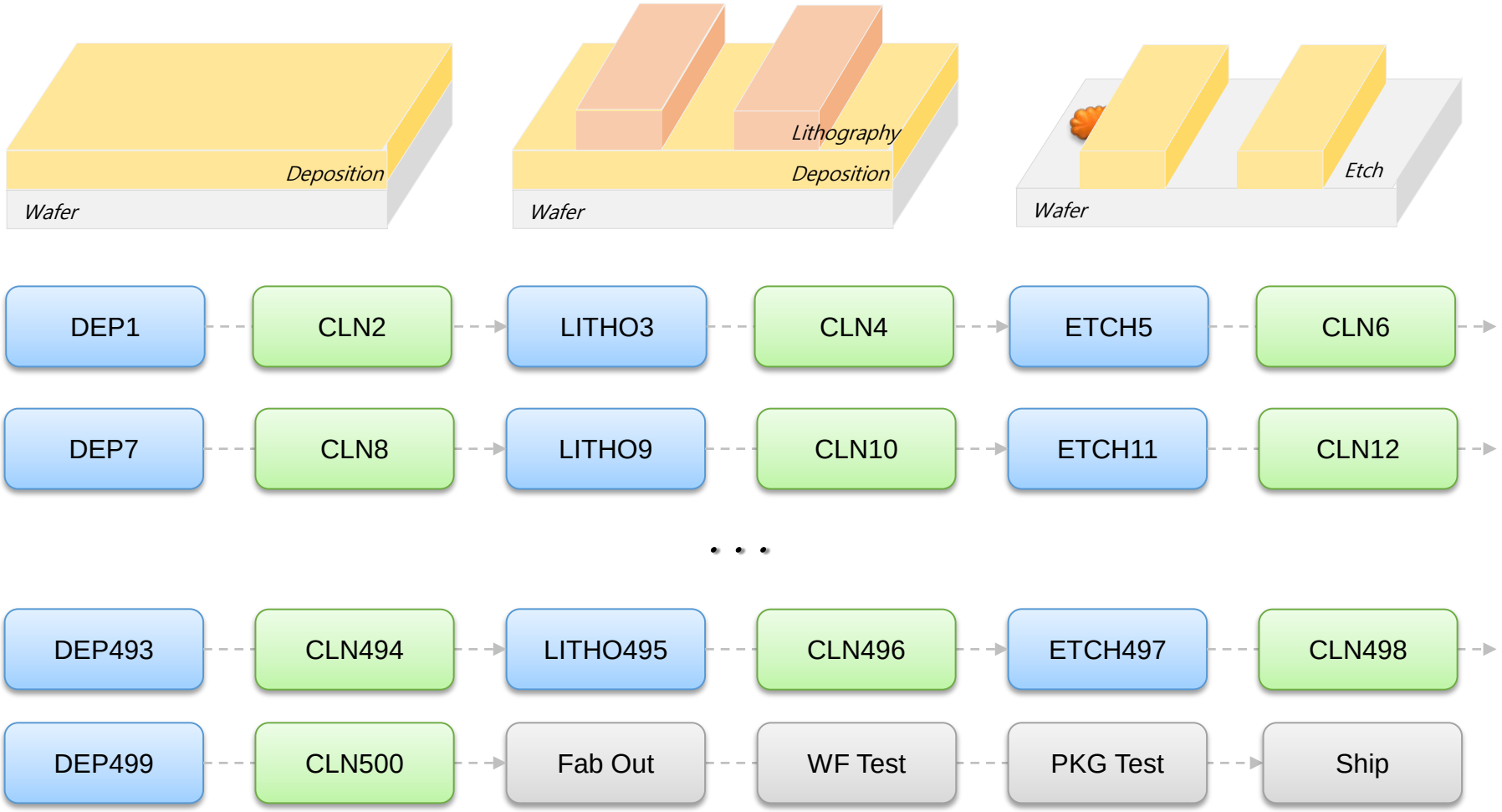
반도체 제조 공정 이해

- **CLN 공정을 진행 하더라도 부산물 남을 가능성 있음**



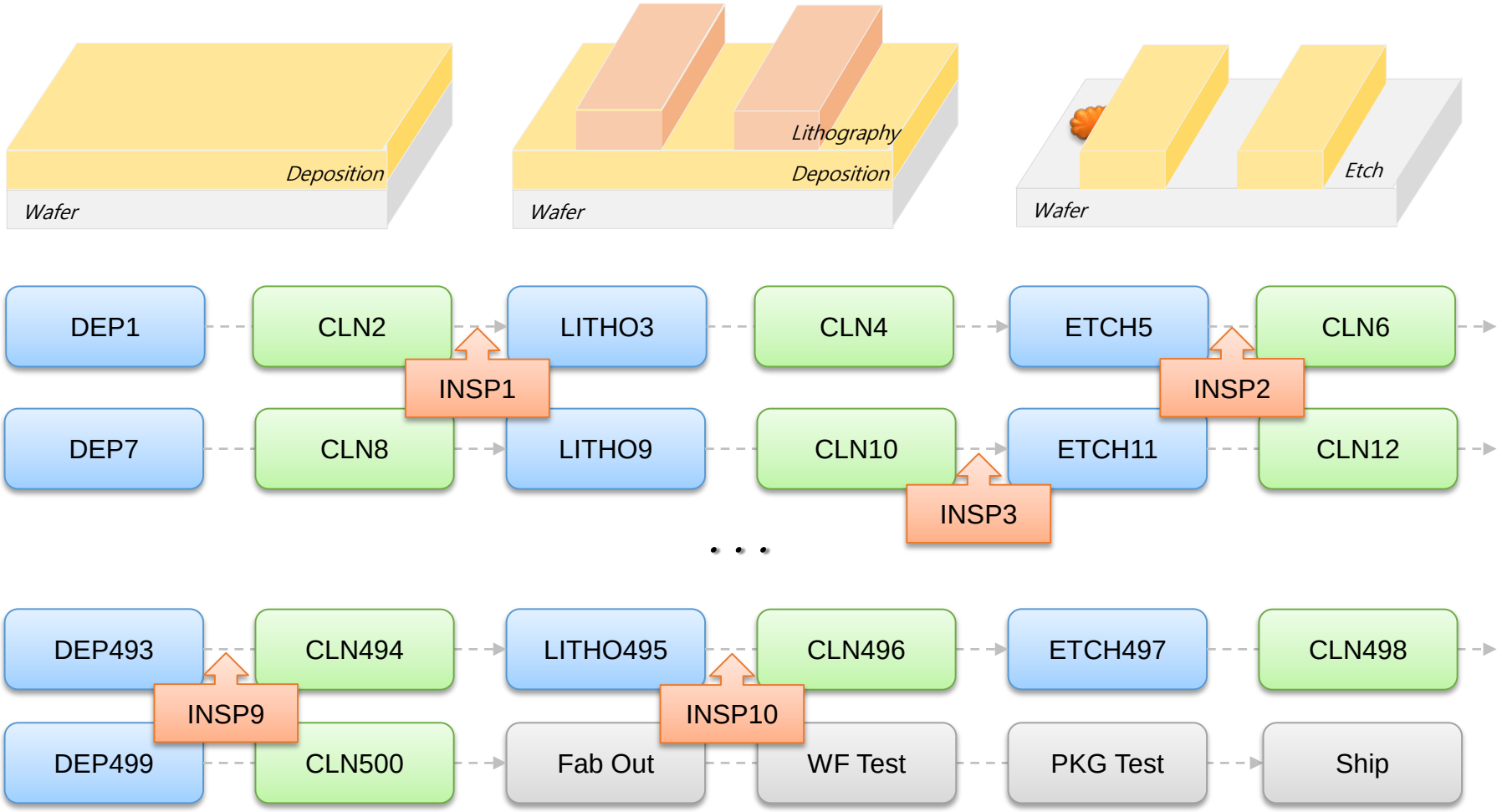
반도체 제조 공정 이해

- 여러 공정 진행 시 어떤 공정 문제인지 파악 어려움



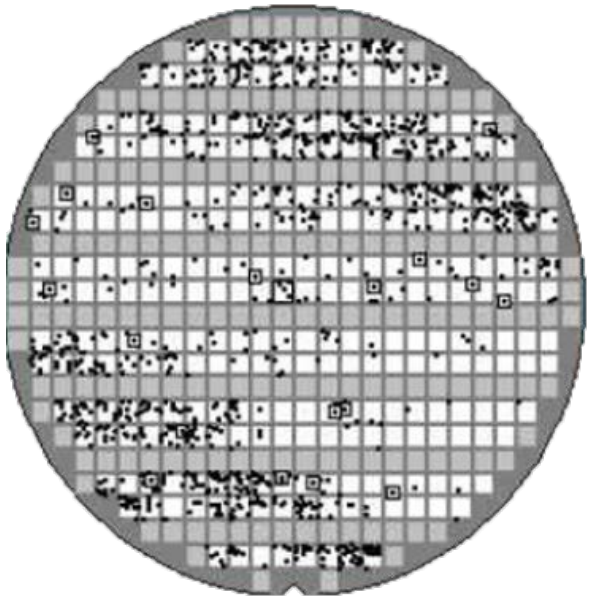
반도체 제조 공정 이해

- 공정 중 Sampling 하여 INSP. 진행 – Defect 발생 여부 확인

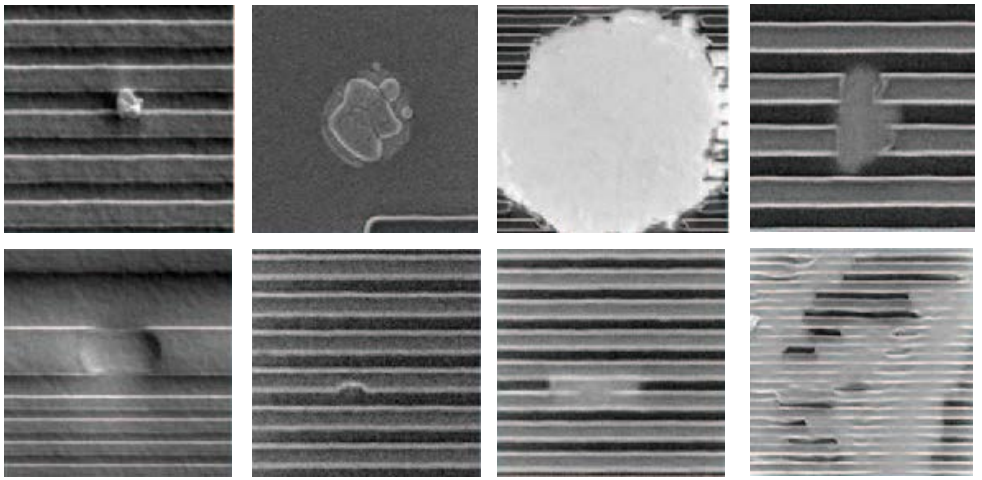


반도체 제조 공정 이해

- **Inspection – Find Abnormalities**

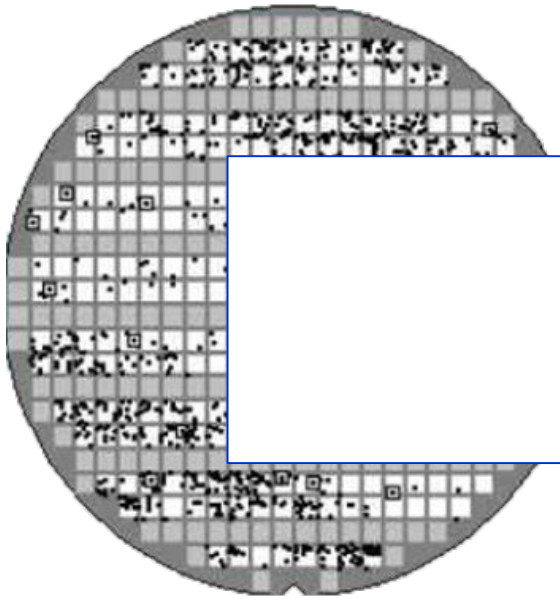


Inspection 장비를 이용하여 Wafer 내 비정상적인 부분 탐색.
탐색 된 IMG 를 바탕으로 해당 공정에 문제가 있는지 확인 가능.
해당 IMG 가 수율과 연관이 있는지 확인 하는 것이 중요.



반도체 제조 공정 이해

- **Inspection – Find Abnormalities**



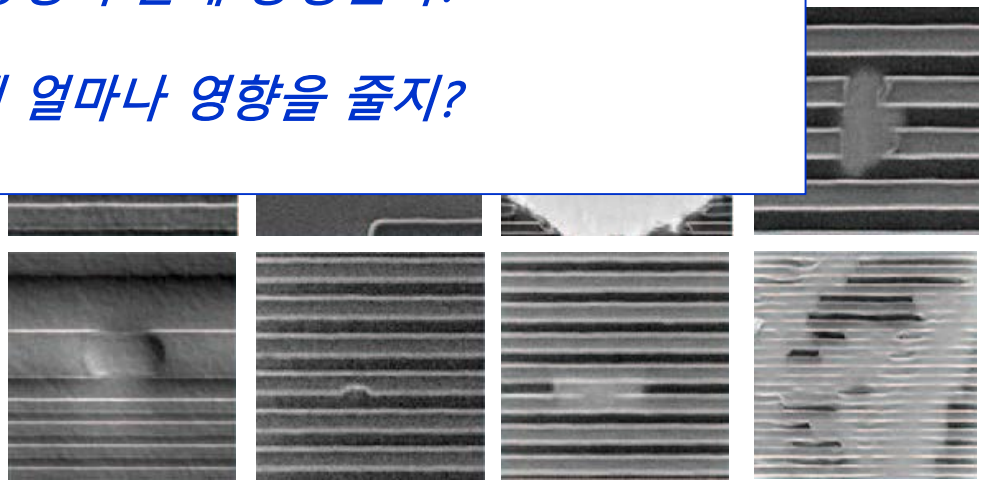
Inspection 장비를 이용하여 Wafer 내 비정상적인 부분 탐색.

탐색 된 IMG 를 바탕으로 해당 공정에 문제가 있는지 확인 가능.

이 중요.

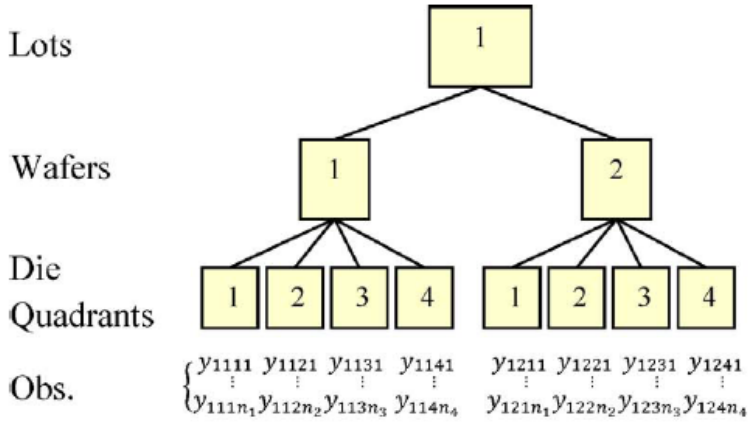
어떤 공정이 문제 공정인지?

수율에 얼마나 영향을 줄지?



Defect data 를 이용한 수율 예측 및 원인 공정 탐지

- Defect data 를 이용한 수율 예측 모델 생성.
- Defect data : particle, scratch, pattern defect와 같은 비정상적인 pattern 의 die 위치, defect 크기, die 내 위치 정보 등을 포함.
- 본 논문에서는 die 내 defect 개수를 이용.
- 기존 연구에서는 die 의 구조를 적절히 고려하지 않고 die 간에 독립을 가정.
  → die – wafer – lot 의 구조. 독립이 아님.



44 IEEE TRANSACTIONS ON SEMICONDUCTOR MANUFACTURING, VOL. 24, NO. 1, FEBRUARY 2011

Application of Generalized Linear Models to Predict Semiconductor Yield Using Defect Metrology Data

Dana C. Krueger, Douglas C. Montgomery, and Christina M. Mastrangelo

***Abstract*—Semiconductor yield modeling is essential to identify processing issues, improve quality, and meet customer demand. However, the massive amounts of data collected during the fabrication process and the number of historical models available make yield modeling a complex and challenging task. This paper presents a methodology to guide the practitioner in determining what data should be collected, integrated, and aggregated, followed by a modeling strategy to forecast yield using generalized linear models based on defect metrology data. This technique yields results at both the die and the wafer levels, significantly outperforms existing models found in the literature based on prediction errors, and identifies significant factors that can drive process improvement. This method also allows the nested structure of the process to be considered in the model, improving predictive capabilities and violating fewer assumptions. An example is presented to discuss this approach and to demonstrate the advantages of these models over the models of the past.**

***Index Terms*—Defect modeling, generalized linear models, logistic regression, semiconductor yield, yield modeling.**

I. Introduction

YIELD IS A key process performance characteristic in the capital-intensive semiconductor fabrication process. Semiconductor yield may be defined as the fraction of total input transformed into shippable output [1]. In an industry where machines cost millions of dollars and cycle times are a number of months, predicting and optimizing yield are critical to process improvement, customer satisfaction, and financial success.

Since the 1960s, semiconductor yield models have been used in the planning, optimization, and control of the fabrication process [2]. A comprehensive review of these methods is given in [3]. Many of these methods focus on using defect metrology information, sometimes referred to as *defectivity* data, to predict yield. While several other measurements, such as critical dimensions and electrical tests, are taken as wafers are fabricated, defectivity data seem to be the most influential in current yield modeling practice.

Defectivity measures come from a wafer-surface scan that identifies unusual patterns such as particles, scratches, or pattern defects. These scans are performed after different layers of the wafer have completed processing. The scans are time consuming, so only a few wafers are sampled to monitor the process and to predict yield. Several types of data may be recorded for each defect, including the die the defect appears on, the size of the defect, and the location of the defect on the die. In addition, a sample of the defects is often selected for classification based on scanning electron microscope images.

Another type of test is performed at *wafer sort*. At this stage, the wafers have completed the fabrication process, and each die on the wafer is tested for functionality. Dice that pass this test move on to be assembled and packaged before a *final test* is performed and the good product is shipped to the customer. At wafer sort, the dice are grouped into *bins*. Passing dice are placed into one bin, while failing dice are separated by their failure modes into a number of different bins.

One of the challenges of working with semiconductor measurements is the size of the massive datasets available with computer-aided manufacturing. While these data record many important process parameters and test results, integrating them into a usable form is a considerable problem. Often, process data are stored in one database, defectivity data in another database, and electrical and wafer sort data in yet a third database. Obtaining a dataset that contains defectivity data and the corresponding wafer sort data can require skilled knowledge of two different systems and the ability to query in both. Aggregating the data into a more useable form for model building is also a time-consuming task.

Another challenge is developing an adequate yield model. Yield models in the literature that use defect metrology data have neglected to properly account for the nested structure of the data and have assumed independence among the data. Dice are grouped together on wafers, and wafers are processed together as lots, making this assumption questionable at best. The yield models in the literature have overlooked this potential source of variation. Also, most current modeling is done at the wafer level, which loses the vast amounts of information available at the die level. In industry, many companies develop their own proprietary yield models that are not available in published literature. Some of the most common methodologies

Manuscript received July 22, 2009; revised August 21, 2010; accepted August 23, 2010. Date of publication October 28, 2010; date of current version February 4, 2011.

D. C. Krueger is with the Department of Management, Kansas State University, Manhattan, KS 66506 USA (e-mail: dkrueger@ksu.edu).
D. C. Montgomery is with the Department of Industrial Engineering, Arizona State University, Tempe, AZ 85287 USA (e-mail: doug.montgomery@asu.edu).
C. M. Mastrangelo is with the Department of Industrial Engineering, University of Washington, Seattle, WA 98195 USA (e-mail: mastr@u.washington.edu).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.
Digital Object Identifier 10.1109/TSM.2010.2089377

Defect data 를 이용한 수율 예측 및 원인 공정 탐지

Data Description

- 18 lots, 36 wafers, 4413 dies (Defect 이 있는 Die 만 고려)
- Training set 12 lots, Test set 6 lots
- 각각의 Die 에 있는 Layer 별 Defects 개수를 이용.

Lot Number	Wafer ID	Die X	Die Y	Radial Distance	Die Quadrant	Layer 1	Layer 2	Layer 3	Layer 4	Layer 5	Layer 6	Layer 7	Layer 8	Layer 9	Layer 10	Response (Fail = 1)
1	2	2	7	7.28	3	1	0	0	0	0	0	0	0	0	0	0
1	2	2	10	7.07	2	0	0	0	0	0	1	0	0	0	0	0
1	2	2	11	7.28	2	0	0	1	0	0	0	0	0	0	0	0
1	2	2	12	7.62	2	0	0	1	0	0	0	0	0	0	0	0
1	2	3	12	6.71	2	0	1	0	0	0	0	0	0	0	0	0
1	2	4	3	7.81	3	1	0	0	0	0	0	0	0	0	0	0
1	2	5	6	5.00	3	0	1	0	0	0	0	0	0	0	0	0
1	2	5	7	4.47	3	1	0	0	0	0	0	0	0	0	0	0
1	2	5	11	4.47	2	1	0	0	0	0	0	0	0	0	0	0
1	2	5	15	7.21	2	1	0	0	0	0	0	0	0	0	0	0
1	2	6	3	6.71	3	0	0	1	0	0	0	0	0	0	0	0
1	2	6	15	6.71	2	1	0	0	0	0	0	0	0	0	0	0
1	2	7	2	7.28	3	0	0	0	1	0	0	0	0	0	0	0
1	2	7	3	6.32	3	1	0	0	0	0	0	0	0	0	0	0
1	2	7	5	4.47	3	0	0	1	0	0	0	0	0	0	0	0
1	2	7	7	2.83	3	0	0	0	0	1	0	0	0	0	0	1
1	2	7	11	2.83	2	0	1	0	0	0	0	0	0	0	0	0

Defect data 를 이용한 수율 예측 및 원인 공정 탐지

Data Description

- 18 lots, 36 wafers, 4413 dies (Defect 이 있는 Die 만 고려)
- Training set 12 lots, Test set 6 lots
- 각각의 Die 에 있는 Layer 별 Defects 개수를 이용.

Lot Number	Wafer ID	Die X	Die Y	Radial Distance	Die Quadrant	Layer 1	Layer 2	Layer 3	Layer 4	Layer 5	Layer 6	Layer 7	Layer 8	Layer 9	Layer 10	Response (Fail = 1)
1	2	2	7	7.28	3	1	0	0	0	0	0	0	0	0	0	0
1	2	2	10	7.07	2	0	0	0	0	0	1	0	0	0	0	0
1	2	2	11	7.28	2	0	0	1	0	0	0	0	0	0	0	0
1	2	2	12	7.62	2	0	0	1	0	0	0	0	0	0	0	0
1	2	3	12	6.71	2	0	1	0	0	0	0	0	0	0	0	0
1	2	4	3	7.81	3	1	0	0	0	0	0	0	0	0	0	0
1	2	5	6	5.00	3	0	1	0	0	0	0	0	0	0	0	0
1	2	5	7	4.47	3	1	0	0	0	0	0	0	0	0	0	0
1	2	5	11	4.47	2	1	0	0	0	0	0	0	0	0	0	0
1	2	5	15	7.21	2	1	0	0	0	0	0	0	0	0	0	0
1	2	6	3	6.71	3	0	0	1	0	0	0	0	0	0	0	0
1	2	6	15	6.71	2	1	0	0	0	0	0	0	0	0	0	0
1	2	7	2	7.28	3	0	0	0	1	0	0	0	0	0	0	0
1	2	7	3	6.32	3	1	0	0	0	0	0	0	0	0	0	0
1	2	7	5	4.47	3	0	0	1	0	0	0	0	0	0	0	0
1	2	7	7	2.83	3	0	0	0	0	1	0	0	0	0	0	1
1	2	7	11	2.83	2	0	1	0	0	0	0	0	0	0	0	0

1. 어떤 Layer 가 중요 변수인지?

Defect data 를 이용한 수율 예측 및 원인 공정 탐지

Data Description

- 18 lots, 36 wafers, 4413 dies (Defect 이 있는 Die 만 고려)
- Training set 12 lots, Test set 6 lots
- 각각의 Die 에 있는 Layer 별 Defects 개수를 이용.

Lot Number	Wafer ID	Die X	Die Y	Radial Distance	Die Quadrant	Layer 1	Layer 2	Layer 3	Layer 4	Layer 5	Layer 6	Layer 7	Layer 8	Layer 9	Layer 10	Response (Fail = 1)
1	2	2	7	7.28	3	1	0	0	0	0	0	0	0	0	0	0
1	2	2	10	7.07	2	0	0	0	0	0	1	0	0	0	0	0
1	2	2	11	7.28	2	0	0	1	0	0	0	0	0	0	0	0
1	2	2	12	7.62	2	0	0	1	0	0	0	0	0	0	0	0
1	2	3	12	6.71	2	0	1	0	0	0	0	0	0	0	0	0
1	2	4	3	7.81	3	1	0	0	0	0	0	0	0	0	0	0
1	2	5	6	5.00	3	0	1	0	0	0	0	0	0	0	0	0
1	2	5	7	4.47	3	1	0	0	0	0	0	0	0	0	0	0
1	2	5	11	4.47	2	1	0	0	0	0	0	0	0	0	0	0
1	2	5	15	7.21	2	1	0	0	0	0	0	0	0	0	0	0
1	2	6	3	6.71	3	0	0	1	0	0	0	0	0	0	0	0
1	2	6	15	6.71	2	1	0	0	0	0	0	0	0	0	0	0
1	2	7	2	7.28	3	0	0	0	1	0	0	0	0	0	0	0
1	2	7	3	6.32	3	1	0	0	0	0	0	0	0	0	0	0
1	2	7	5	4.47	3	0	0	1	0	0	0	0	0	0	0	0
1	2	7	7	2.83	3	0	0	0	0	1	0	0	0	0	0	1
1	2	7	11	2.83	2	0	1	0	0	0	0	0	0	0	0	0

2. 수율 예측을 얼마나 잘할지?

Defect data 를 이용한 수율 예측 및 원인 공정 탐지

1. 어떤 Layer 가 중요 변수인지?

- Die-Level & Wafer-Level Logistic Regression Model 평가

Category	Die Level				Wafer Level		
	Non-Nested	Nested			Non-Nested		
	All Training Set Data	All Training Set Data	Outliers Removed (2.5% trimmed)	Outliers Removed (5% trimmed)	All Training Set Data	Outliers Removed (2.5% trimmed)	Outliers Removed (5% trimmed)
	N = 2967	N = 2967	N = 2896	N = 2845	N =24 (2967 Die)	N = 24 (2896 Die)	N = 24 (2845 Die)
Significant features (α = 0.10)	Radial distance **Layers 4–10**	Radial distance **Layers 4–10**	Radial distance **Layers 2–10**	Radial distance **Layers 2–10**	**Layer 8**	**Layer 9**	**Layer 9**
Pearson (P-value)	0.000	0.000	0.005	**0.121**	0.867	0.853	0.877

Layer 1 은 중요하지 않다.

Defect data 를 이용한 수율 예측 및 원인 공정 탐지

1. 어떤 Layer 가 중요 변수인지?

- Die-Level & Wafer-Level Logistic Regression Model 평가

Category	Die Level				Wafer Level		
	Non-Nested	Nested			Non-Nested		
	All Training Set Data	All Training Set Data	Outliers Removed (2.5% trimmed)	Outliers Removed (5% trimmed)	All Training Set Data	Outliers Removed (2.5% trimmed)	Outliers Removed (5% trimmed)
	N = 2967	N = 2967	N = 2896	N = 2845			
Significant features (α = 0.10)	Radial distance **Layers 4–10**	Radial distance **Layers 4–10**	Radial distance **Layers 2–10**	Radial distance **Layers 2–10**			
Pearson (P-value)	0.000	0.000	0.005	**0.121**	0.867	0.853	0.877

$$P = \frac{1}{1 + e^{-\mathbf{x}_i'\boldsymbol{\beta}}} \quad (9)$$

where $\mathbf{x}_i'\boldsymbol{\beta} = -2.30820 + 0.0664864(\text{radial distance}) + 0.456202(\text{Layer 2}) + 0.222156(\text{Layer 3}) + 0.322290(\text{Layer 4}) + 0.322418(\text{Layer 5}) + 0.877724(\text{Layer 6}) + 0.867638(\text{Layer 7}) + 0.553416(\text{Layer 8}) + 0.947011(\text{Layer 9}) + 0.720564(\text{Layer 10})$.

Defect data 를 이용한 수율 예측 및 원인 공정 탐지

2. 수율 예측을 얼마나 잘할지?

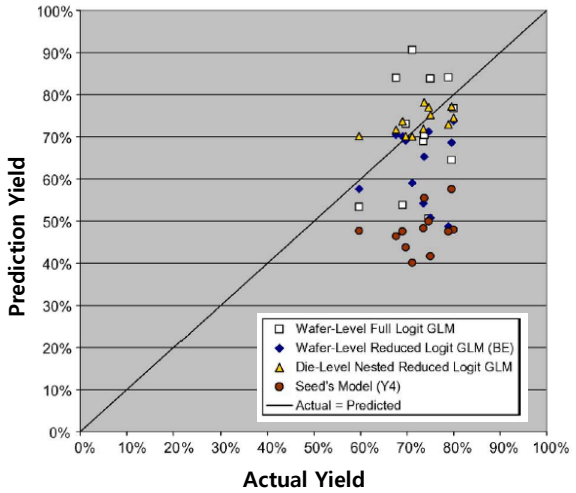
- Die-level & wafer-level model's yield prediction result

Existing yield model

Popular Name	Model
Classic Poisson model:	$Y_1 = e^{-D_0 A}$
Binomial yield model:	$Y_2 = \left[1 - A/A_w\right]^{D_0 A_w}$
Murphy's yield model:	$Y_3 = \left[\frac{1 - e^{-D_0 A}}{D_0 A}\right]^2$
Seeds' yield model:	$Y_4 = \frac{1}{1+D_0 A}$
Dingwall's yield model:	$Y_5 = \left[1 + D_0 A/3\right]^{-3}$
Moore's yield model:	$Y_6 = e^{-\sqrt{D_0 A}}$
Price's yield model:	$Y_7 = \prod_{i=1}^{n} \frac{1}{1+D_i A}$
Price's general model:	$Y_8 = (1 + D_0 A/n)^{-n}$
Negative binomial model:	$Y_9 = (1 + D_0 A/\alpha)^{-\alpha}$

D0 = defects per area
A = area of a die

Predicted vs. Actual Wafer Level Yields



Die-Level Nested Reduced Logit GLM Model 이 가장 우수한 MSE,MAD 값을 나타냄

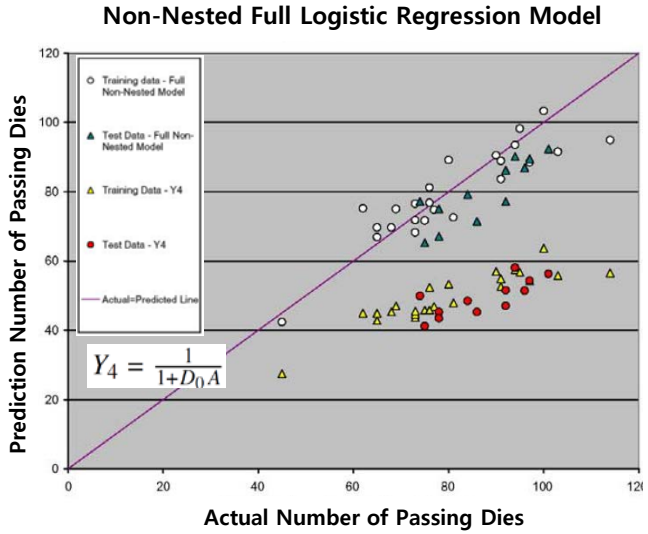
MSE AND MAD FOR MODEL COMPARISONS AT THE WAFER LEVEL (%)

	Y1	Y2	Y3	Y4	Y5	Y6	Y7	Y8	Y9	Wafer-Level Full Logit GLM	Wafer-Level Reduced Logit GLM (BE)	Die-Level Nested Reduced Logit GLM
MAD	0.516	0.516	0.481	0.373	0.457	0.500	0.486	0.496	0.462	0.162	0.219	0.116
MSE	0.269	0.270	0.234	0.141	0.212	0.251	0.239	0.250	0.217	0.037	0.057	0.014

Defect data 를 이용한 수율 예측 및 원인 공정 탐지

2. 수율 예측을 얼마나 잘할지?

- Non-nested & nested model's yield prediction result



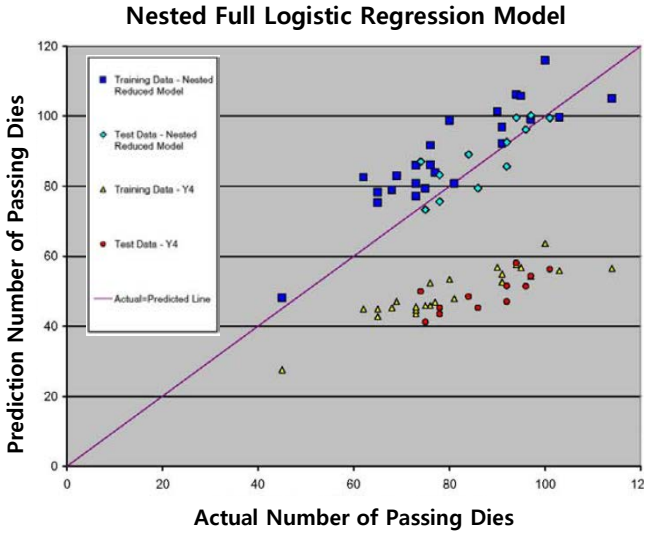


TABLE VII

MSE AND MAD FOR MODEL COMPARISONS AT THE DIE LEVEL USING TEST DATA

	Y1	Y3	Y4	Y5	Y6	Y7	Y8	Y9	Full Nested Logit	Reduced Nested Logit	Full Non-Nested Die-Level Logit	Multiple Linear Regression
MSE	0.412	0.377	0.289	0.358	0.371	0.309	0.394	0.362	0.190	0.189	0.188	0.592
MAD	0.583	0.566	0.517	0.556	0.572	0.522	0.574	0.558	0.356	0.356	0.397	0.650

Defect data 를 이용한 수율 예측 및 원인 공정 탐지

Conclusions

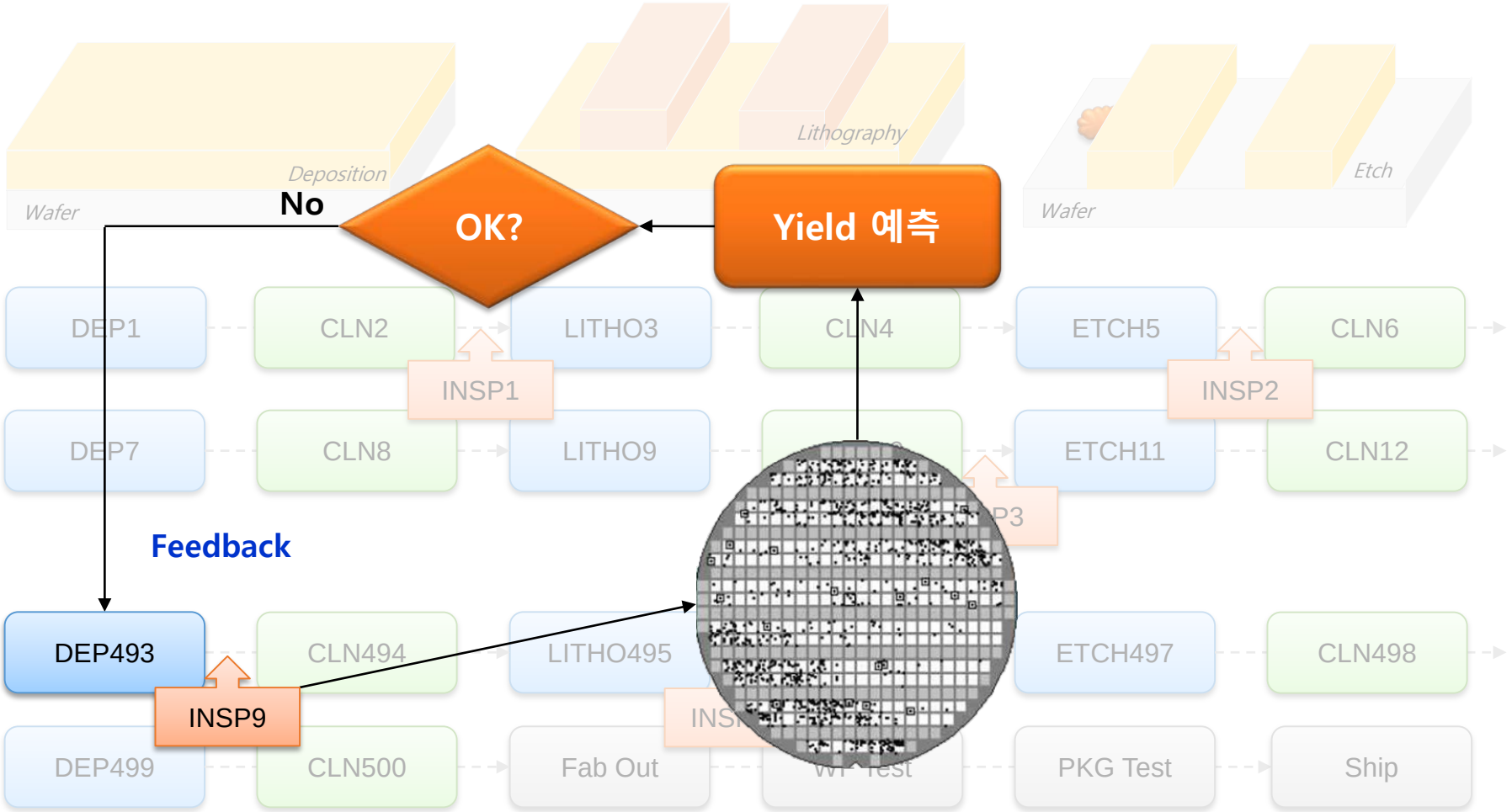
- Logistic Regression 을 사용하여 Defect Data 를 기반으로 한 반도체 수율을 성공적으로 모델링 함.
- 기존 Wafer Level Model 대비 Die Level Model 에서 Yield 를 잘 예측 함.
- Nested 구조를 사용함으로써 더 좋은 MSE,MAD 값을 확보 함.
- 중요한 예측 변수 식별함. 문제 진단 및 개선이 필요한 영역을 찾는데 도움이 됨.

Limitations

- Defect 유형, 크기, Die 내 위치 정보를 이용한다면 예측력이 더 향상 될 수 있다.
- 각 변수의 중요도를 제공할 수 있다면 가장 먼저 확인 해야할 공정을 더 쉽게 찾을 수 있다.
- Nested 구조 구축 시 Wafer Quad. 정보를 이용하였는데 이보다 공정에서 주로 발생 되고 있는 Center-Middle-Edge 로 구축을 하였다면 더 좋은 결과를 얻을 수 있다.
- Defect Data 정보를 단순히 Count 를 이용하는 것이 아니라, 주변의 Defect 상황을 판단 하여 어떤 유형의 Pattern 을 가지는 Defect 인지에 대한 정보를 고려한다면 더 좋은 예측 모델이 될 수 있다.

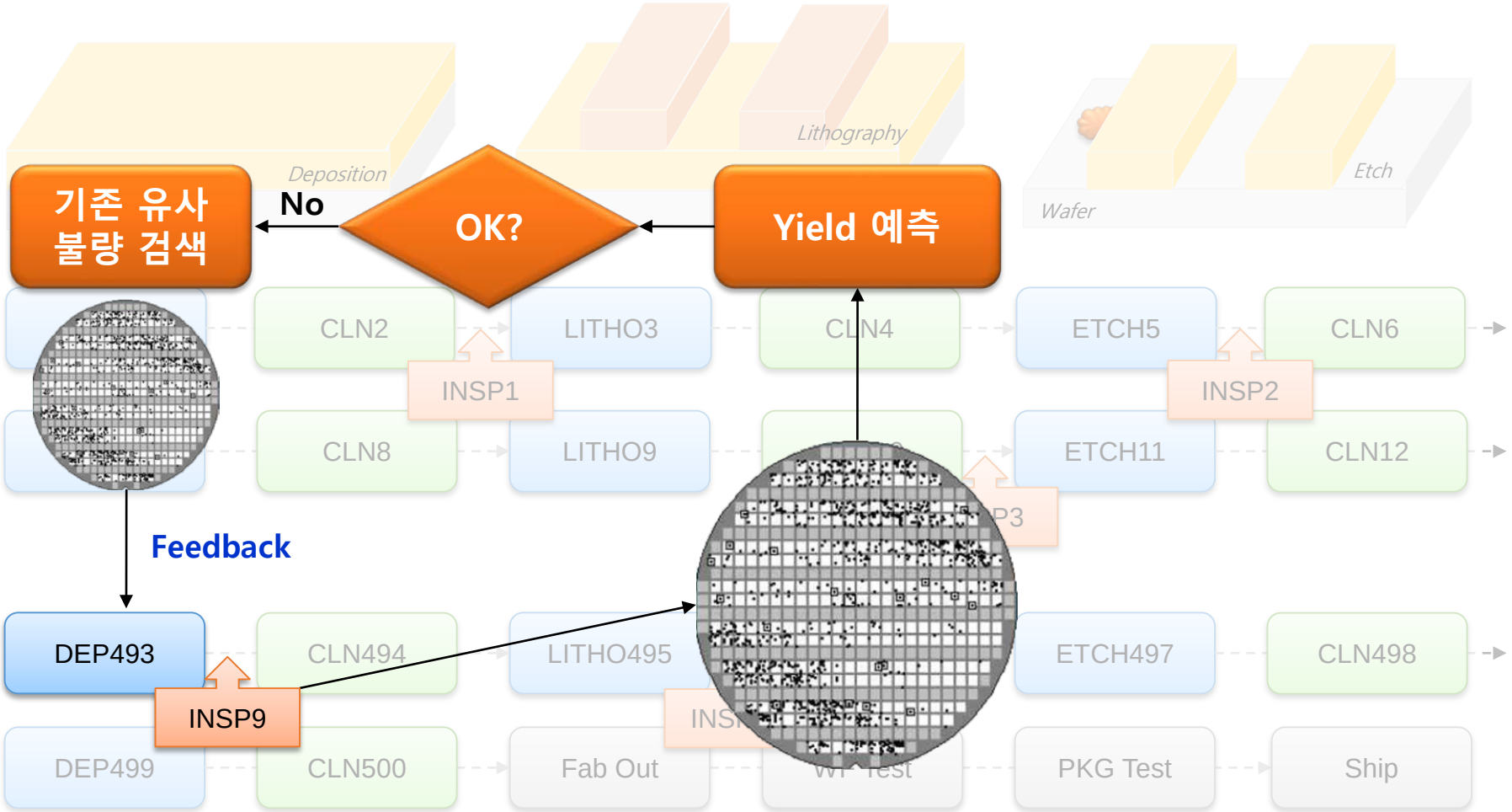
Defect data 를 이용한 수율 예측 및 원인 공정 탐지

- **Yield 예측을 통한 문제 상황 확인 및 공정 Feedback**



Defect data 를 이용한 수율 예측 및 원인 공정 탐지

- **Yield 예측을 통한 문제 상황 확인 및 공정 Feedback**



Defect pattern 을 이용한 Wafer 분류 및 유사 Wafer 검색

- Wafer map 은 defect pattern 을 시각화 하고 잠재적 공정 문제를 식별 하는데 사용
- 생성 된 wafer map 을 분류하거나 wafer map library 에서 검색하여 유사한 wafer 를 찾고 어떤 문제가 있었는지 확인을 통하여 공정 문제에 신속하게 대응 가능.
- Defect pattern classification 과 Image retrieval 을 위하여 CNN Model 구축.

Wafer failure mode classification

Similar failure mode wafer retrieval

IEEE TRANSACTIONS ON SEMICONDUCTOR MANUFACTURING, VOL. 31, NO. 2, MAY 2018 309

Wafer Map Defect Pattern Classification and Image Retrieval Using Convolutional Neural Network

Takeshi Nakazawa and Deepak V. Kulkarni

***Abstract*—Wafer maps provide important information for engineers in identifying root causes of die failures during semiconductor manufacturing processes. We present a method for wafer map defect pattern classification and image retrieval using convolutional neural networks (CNNs). Twenty eight thousand six hundred synthetic wafer maps for 22 defect classes are generated theoretically and used for CNN training, validation, and testing. The overall classification accuracy for the 6600 test dataset is 98.2%. One thousand one hundred and ninety one real wafer maps are used for CNN performance evaluation for the same model trained by synthetic wafer maps. We demonstrate that by using only synthetic data for network training, real wafer maps can be classified with high accuracy. For image retrieval, a binary code for each wafer map is generated from an output of a fully connected layer with sigmoid activation. A retrieval error rate is 0.36% for the test dataset and 3.7% for the real wafers. Image retrieval takes 0.13 s per wafer map from the 18 000 wafer map library.**

***Index Terms*—Deep learning, convolutional neural network, information retrieval, semiconductor defects.**

I. Introduction

IN THE semiconductor manufacturing, wafer maps are used to visualize defect patterns and identify potential process issues. Inline metrology tools perform inspection after a certain process step and monitor abnormalities on dies. Then a wafer map is created based on the detected abnormal locations. One of the main purposes for wafer map visualization is to monitor any abnormal defect signatures and respond to process problems quickly. Once wafer map libraries are created with corresponding root causes, defect pattern similarities between wafers could be a good indication of the common root causes and this knowledge base can be used to solve problems. In order to have an effective knowledge base, two components are required: 1) wafer map defect pattern classification and 2) wafer map image retrieval from historical wafer map libraries. The wafer map defect pattern classification can provide information about a defect occurrence rate for each defect class and engineers focus on the most important issue using this data. The wafer map image retrieval is helpful to identify a root cause by querying historical wafer maps with the known root cause.

Manuscript received November 2, 2017; revised December 31, 2017; accepted January 15, 2018. Date of publication January 18, 2018; date of current version May 8, 2018. *(Corresponding author: Takeshi Nakazawa.)*

The authors are with Intel Corporation, Chandler, AZ 85226 USA (e-mail: takeshi.nakazawa@intel.com).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TSM.2018.2795466

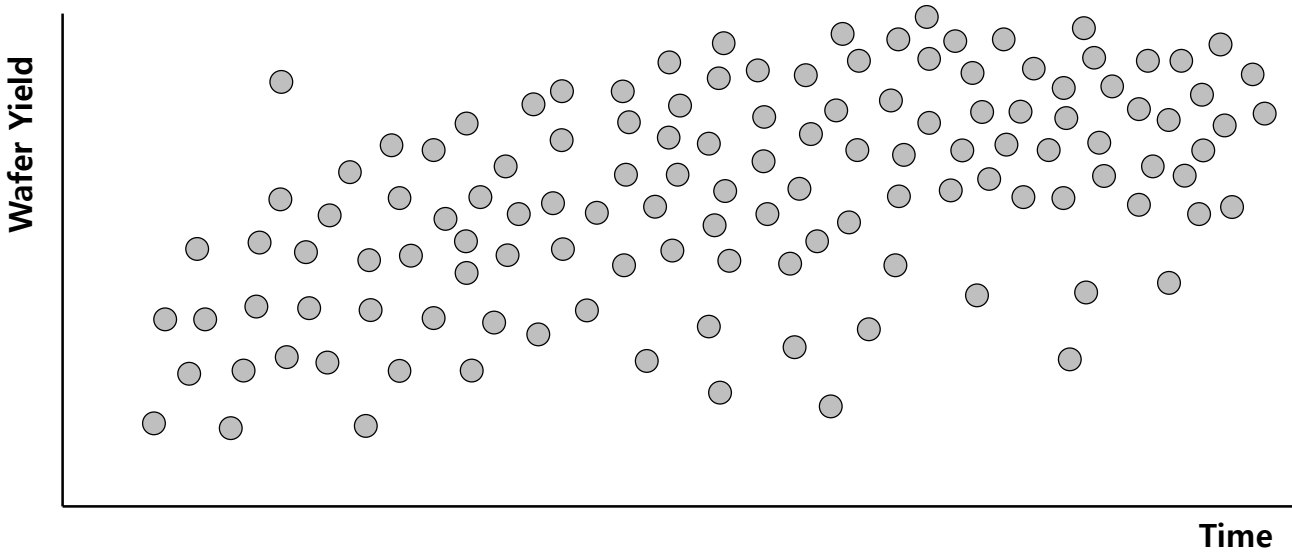
There are a number of studies for wafer map pattern recognitions [1]–[4]. Their classification approaches can be divided into two main groups: 1) model-based pattern recognition, 2) feature extraction based pattern recognition. The model-based pattern recognition uses a predefined probability distribution function for each defect pattern and selects the best matching model using information criterion such as the Akaike information criterion (AIC) and the Bayesian information criterion (BIC). The feature extraction based pattern recognition extracts pattern features using techniques such as correlogram and Radon transform. Once the pattern features are extracted, the common pattern classification algorithms such as support vector machines, neural networks, nearest neighbors etc. are applied for the classification task.

Deep convolutional neural networks (CNN) [5] have recently advanced the state-of-the-art image classification performance and became the standard approach for any image classification tasks. CNN is the end-to-end model and does not require any task-specific feature engineering. This end-to-end model approach is beneficial since we don't need to develop the task specific feature extractors and the domain specific export knowledge is not required. Another aspect of image classification is the problem of image retrieval [6], [7]. The image retrieval is a task of finding images containing similar objects or scene, given a query image, and has been used in security and surveillance, medical imaging, and many other areas. Traditionally, the image retrieval requires feature extraction using object color and shapes. Since the deep CNN can learn rich features at each layer, these intermediate features are used as good descriptors for image retrieval [8], [9].

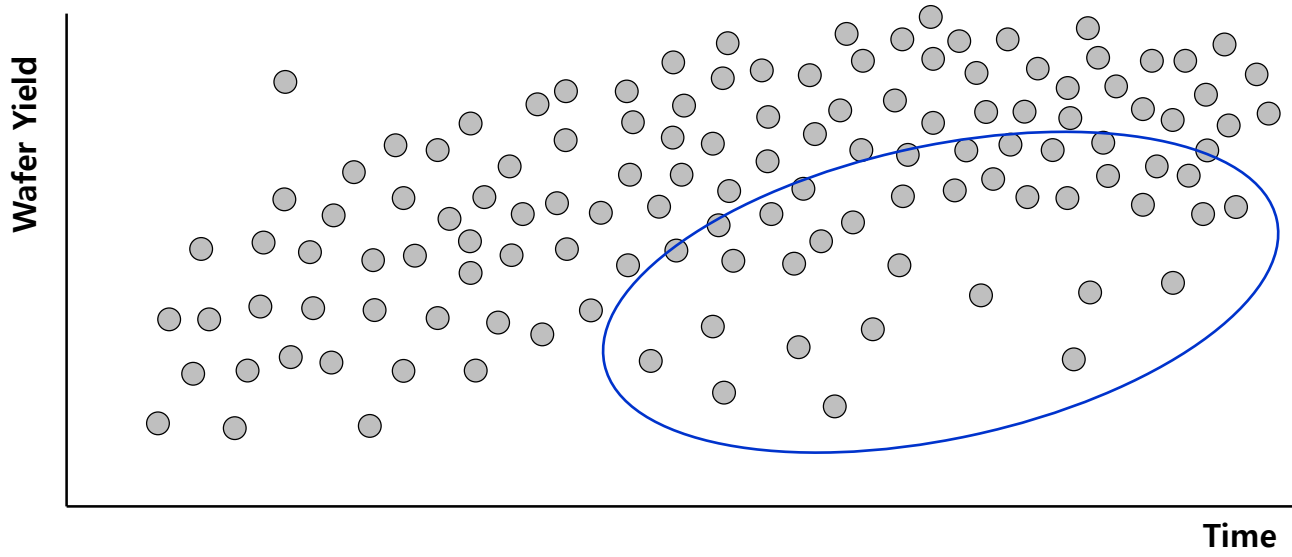
In this paper, we employ CNN for the defect pattern classification and wafer map retrieval tasks. As a dataset, we use wafer maps from simulation and the real wafers. For CNN training and validation, we only use the simulated wafer maps because real data available for each class from the manufacturing process is highly imbalanced. In this case, it is beneficial to train CNN by using theoretically generated data so that we can also include rare defect patterns to the model and yet achieve reasonable classification accuracy. To verify the performance of the proposed method, we generated 28,600 dataset by simulation. Data from 1,191 real wafers are also used to evaluate the performance of the trained CNN.

Our paper is organized as follows. In Section II, methods for wafer map pattern generation, the CNN configuration and CNN based image retrieval are descried. In Section III, we present the results of defective wafer map pattern generations, the CNN training/validation/test results using theoretically

Defect pattern 을 이용한 Wafer 분류 및 유사 Wafer 검색

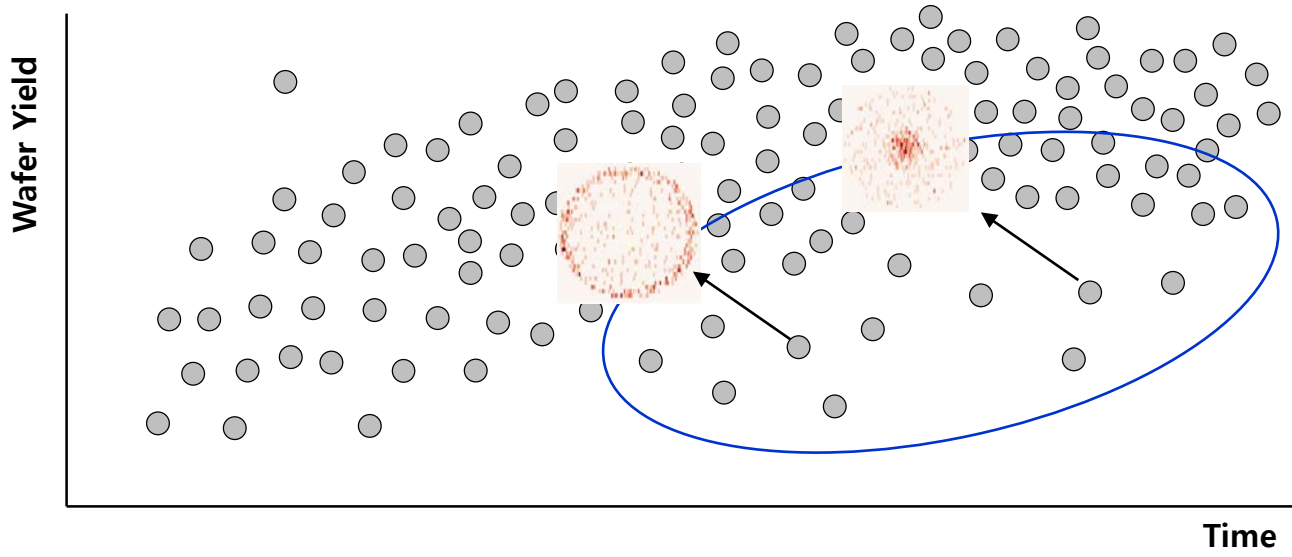


Defect pattern 을 이용한 Wafer 분류 및 유사 Wafer 검색



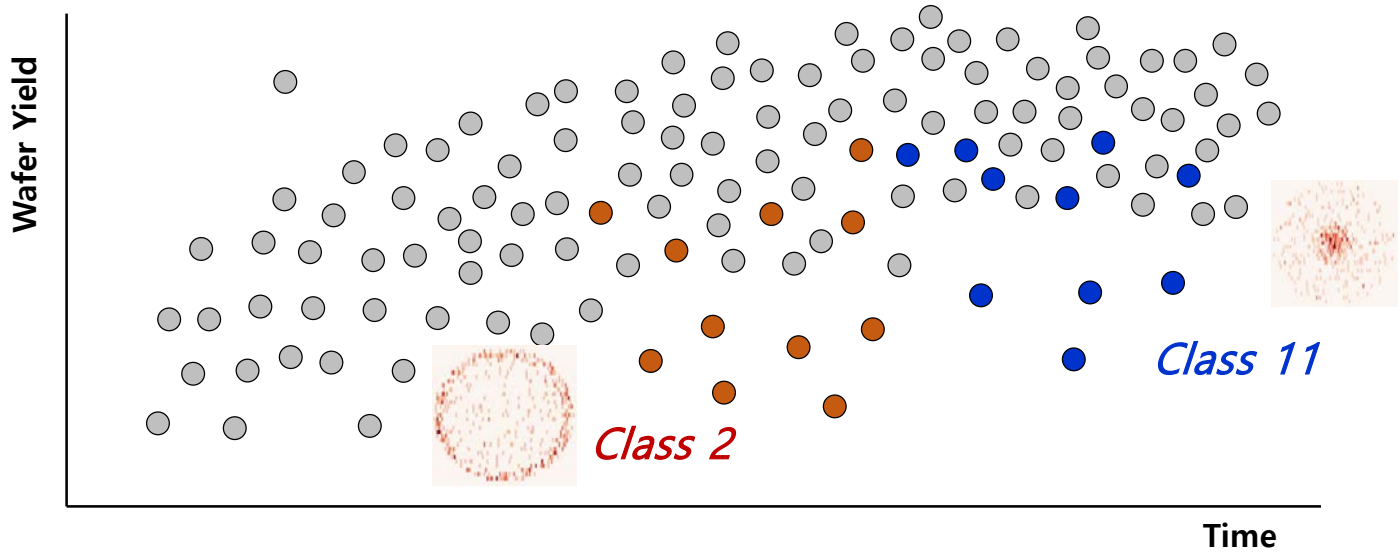
수율 향상을 위해서
낮은 수율 자재의 원인을 파악하고
해당 공정을 개선 시키는 것이 필요.

Defect pattern 을 이용한 Wafer 분류 및 유사 Wafer 검색



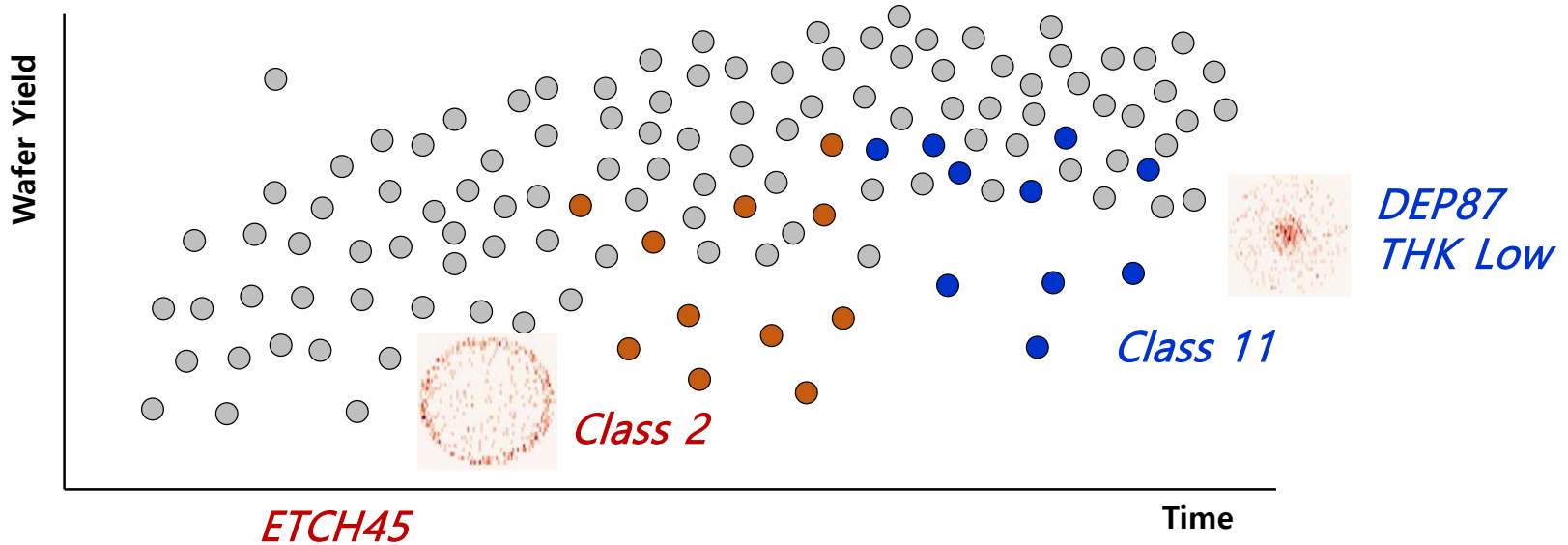
원인 파악을 위하여 Wafer 별 특이 사항을 Check.

Defect pattern 을 이용한 Wafer 분류 및 유사 Wafer 검색



CNN 을 이용하여 Wafer Map 을 분류

Defect pattern 을 이용한 Wafer 분류 및 유사 Wafer 검색

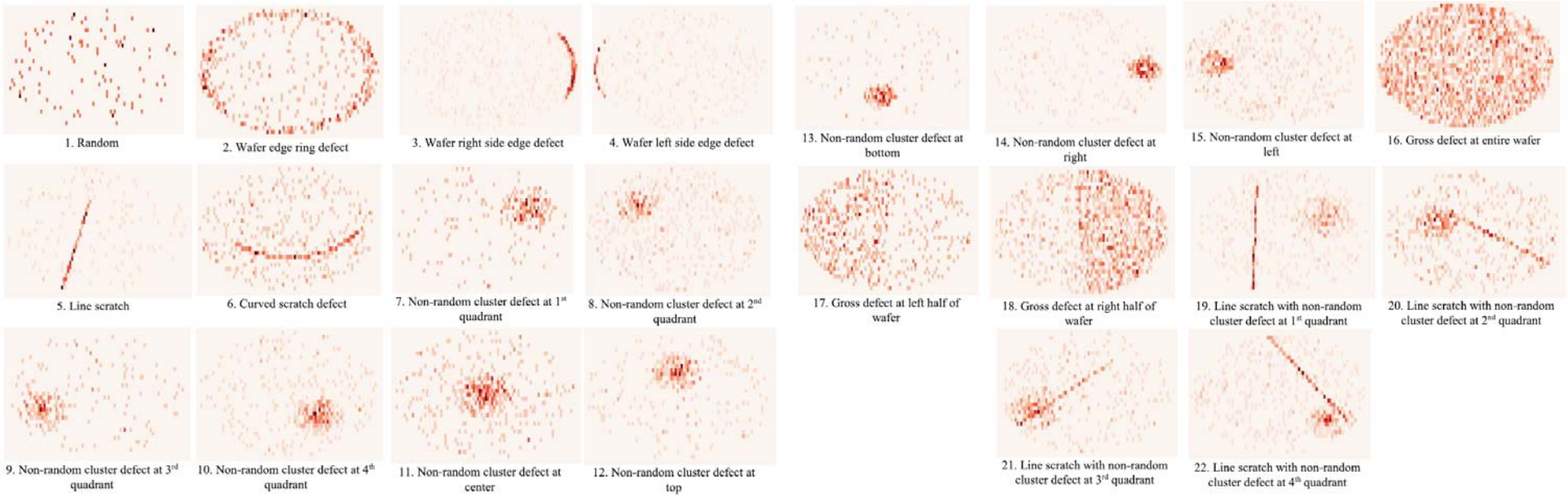


분류된 Map 을 기준으로 검색하여
기존에 유사한 불량 Map 이 어떤 공정
문제였는지, 어떻게 개선 했는지 확인.

Defect pattern 을 이용한 Wafer 분류 및 유사 Wafer 검색

Data Description

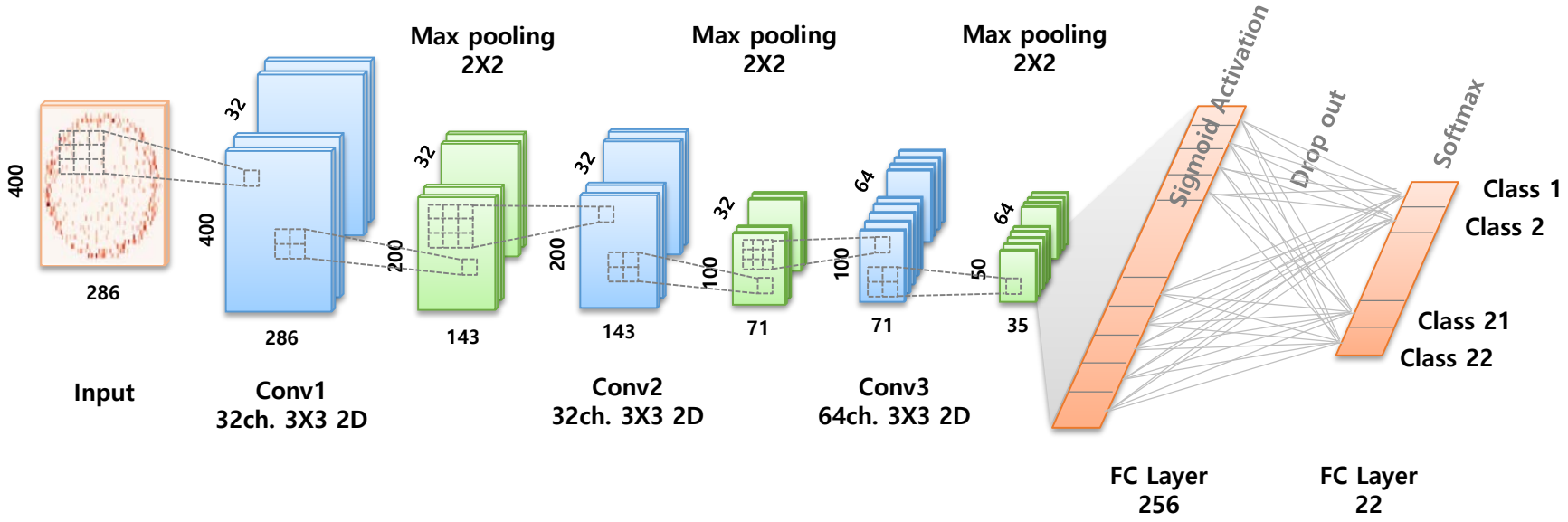
- Training set 28,600 WFs (simulated), Test Set 1,191 WFs (real)
- 아래와 같은 22가지의 class 를 정의하고 해당 class 별로 균등하게(1,300개씩) 유사 이미지를 생성.
- Real data 의 경우 특정 class 의 small data 문제가 있을 수 있고 이런 경우 CNN 의 학습력이 감소 될 가능성이 있으므로 학습할 때는 simulation 으로 생성된 wafer map 을 이용하는 것이 유리.



Defect pattern 을 이용한 Wafer 분류 및 유사 Wafer 검색

Convolutional Neural Network Configuration

- Three convolutional layers with the receptive field size of 3 x 3 and stride 1.
- The rectified linear activation is used for each convolutional layer. The max pooling size is 2 x 2.
- The fully connected (FC) layer with the size of 256 is added after the convolutional layers with sigmoid activation
- The last layer is the softmax layer for the class probability calculation.



Defect pattern 을 이용한 Wafer 분류 및 유사 Wafer 검색

Wafer Map Classification Accuracy

- 각 class 별 1300개의 Simulation map 생성. Train 700개, Validation 300개, Test 300개씩 구분 적용
- 15,400 training, 6,600 validation, 6,600 test 후 높은 accuracy 확보
  10epoc 후, training accuracy : 99.8%, validation accuracy : 97.7%

Test Set, N=6,600

True class label \ Predicted class label	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	C14	C15	C16	C17	C18	C19	C20	C21	C22
C1	98.7	0.0	0.0	0.0	1.0	0.3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
C2	0.0	100.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
C3	0.0	0.0	100.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
C4	0.0	0.0	0.0	100.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
C5	0.3	0.0	0.0	0.0	89.0	10.3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.3	0.0	0.0
C6	1.7	0.0	0.0	0.0	10.3	87.7	0.3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
C7	0.0	0.0	0.0	0.0	0.0	0.0	100.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
C8	0.0	0.0	0.0	0.0	0.0	0.0	0.0	100.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
C9	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	99.7	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.3	0.0
C10	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	100.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
C11	0.3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	99.7	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
C12	0.3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	99.7	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
C13	0.3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.3	0.0	99.4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
C14	0.3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.3	0.0	0.0	99.3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
C15	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.7	0.0	0.0	0.0	99.3	0.0	0.0	0.0	0.0	0.0	0.0	0.0
C16	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	100.0	0.0	0.0	0.0	0.0	0.0	0.0
C17	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	100.0	0.0	0.0	0.0	0.0	0.0
C18	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	100.0	0.0	0.0	0.0	0.0
C19	0.0	0.0	0.0	0.0	1.7	0.0	2.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	96.3	0.0	0.0	0.0
C20	0.0	0.0	0.0	0.0	0.7	0.0	0.0	1.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	98.3	0.0	0.0
C21	0.0	0.0	0.0	0.0	0.3	0.3	0.0	0.0	1.3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	98.0	0.0
C22	0.0	0.0	0.0	0.0	1.7	0.3	0.0	0.0	0.0	3.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	95.0

Real Set, N=1,191

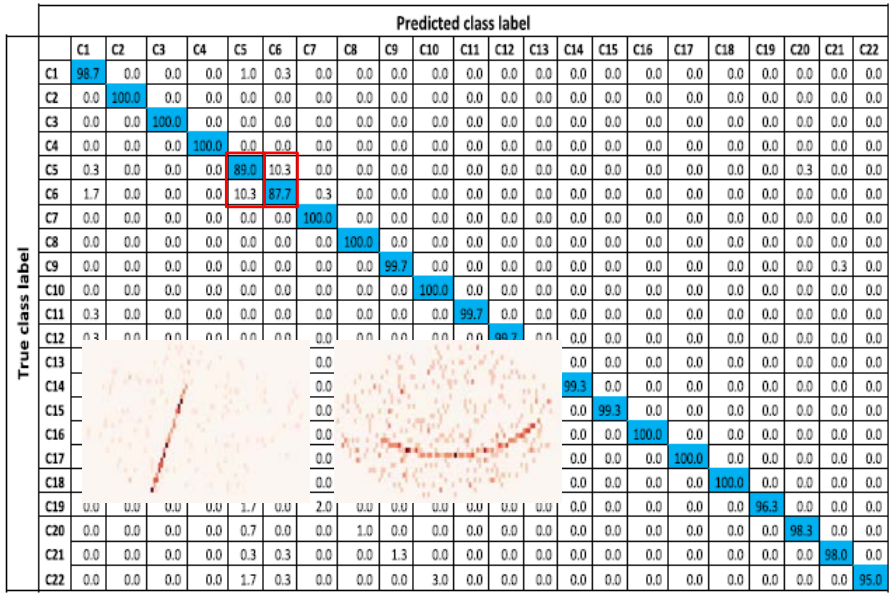
True class label \ Predicted class label	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	C14	C15	C16	C17	C18	C19	C20	C21	C22
C1	99.5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.5	0	0	0	0	0	0
C2	0	100	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
C3	0	0	X	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
C4	0	0	0	X	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
C5	0	0	0	0	X	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
C6	33.3	0	0	0	0	66.7	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
C7	0	0	0	0	0	0	X	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
C8	0	0	0	0	0	0	0	X	0	0	0	0	0	0	0	0	0	0	0	0	0	0
C9	0	0	0	0	0	0	0	0	X	0	0	0	0	0	0	0	0	0	0	0	0	0
C10	0	0	0	0	0	0	0	0	0	X	0	0	0	0	0	0	0	0	0	0	0	0
C11	0	0	0	0	0	0	0	0	0	0	X	0	0	0	0	0	0	0	0	0	0	0
C12	0	0	0	0	0	0	0	0	0	0	0	X	0	0	0	0	0	0	0	0	0	0
C13	0	0	0	0	0	0	0	0	0	0	0	0	X	0	0	0	0	0	0	0	0	0
C14	0	0	0	0	0	0	0	0	0	0	0	0	0	100	0	0	0	0	0	0	0	0
C15	0	0	0	0	0	0	0	0	0	0	0	0	0	0	83.3	0	16.7	0	0	0	0	0
C16	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	100	0	0	0	0	0	0
C17	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	100	0	0	0	0	0
C18	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	100	0	0	0	0
C19	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	X	0	0	0
C20	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	X	0	0
C21	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	X	0
C22	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	100

Defect pattern 을 이용한 Wafer 분류 및 유사 Wafer 검색

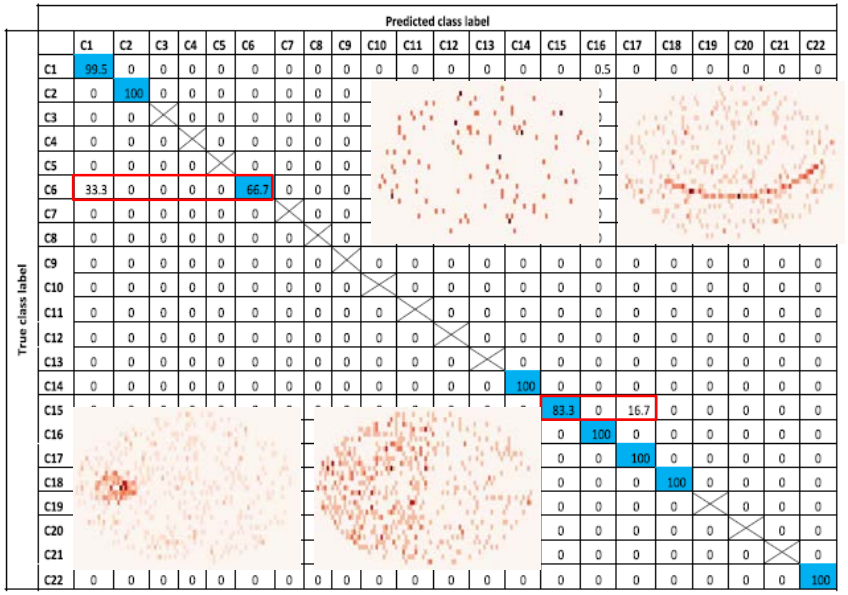
Wafer Map Classification Accuracy

- 각 class 별 1300개의 Simulation map 생성. Train 700개, Validation 300개, Test 300개씩 구분 적용
- 15,400 training, 6,600 validation, 6,600 test 후 높은 accuracy 확보
  10epoc 후, training accuracy : 99.8%, validation accuracy : 97.7%

Test Set, N=6,600



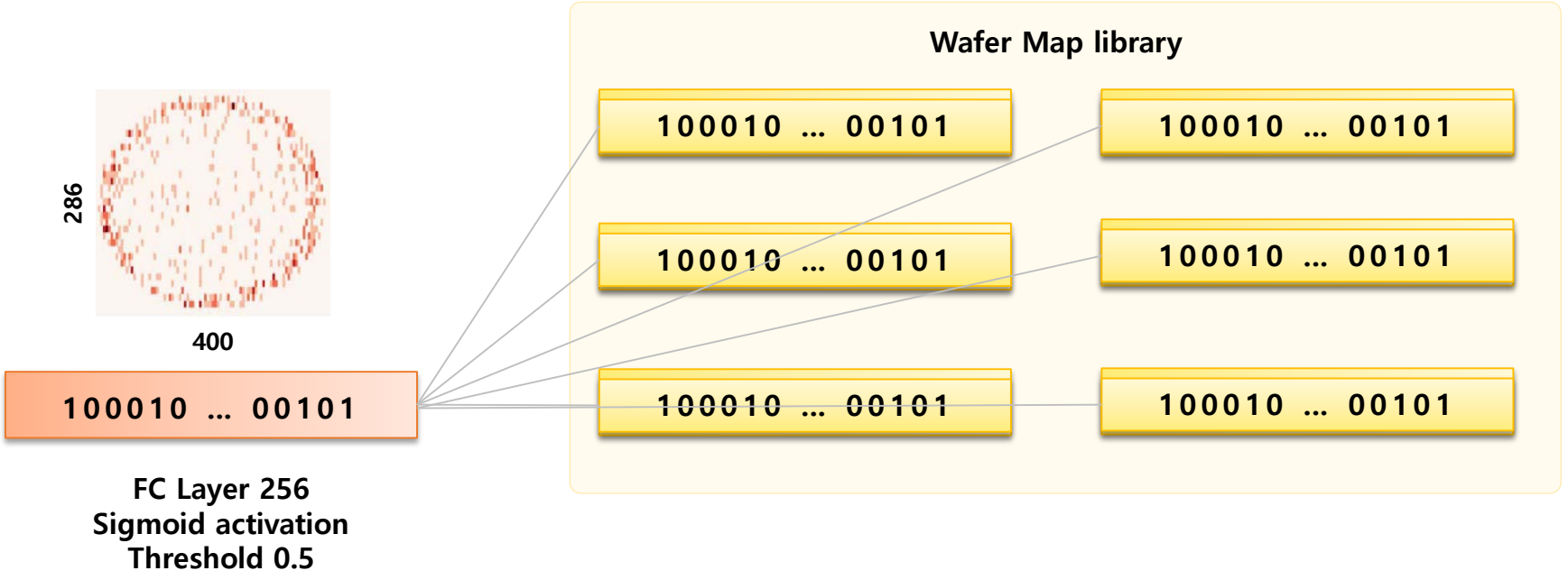
Real Set, N=1,191



Defect pattern 을 이용한 Wafer 분류 및 유사 Wafer 검색

Wafer Map Image Retrieval

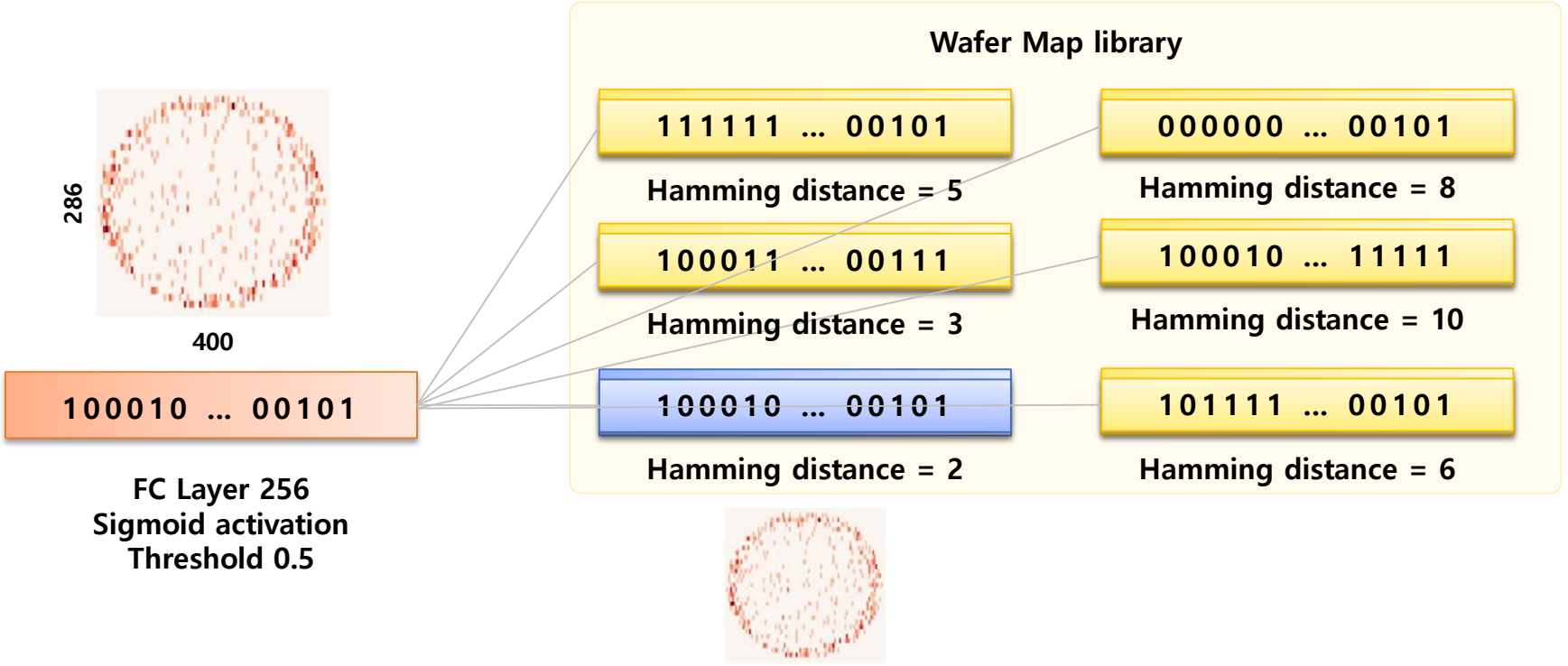
- 이미지는 고차원 데이터이기 때문에 큰 DB에서 빠른 검색을 위해 차원 축소 필수적임.
- 256 Node 의 FC layer 를 이용하여, sigmod activation 후 feature extract 함
  threshold value 0.5로 설정, 0.5 이상이면 1 아니면 0으로 입력
- Hamming distance measure로 주어진 쿼리 wafer map에 적용



Defect pattern 을 이용한 Wafer 분류 및 유사 Wafer 검색

Wafer Map Image Retrieval

- 이미지는 고차원 데이터이기 때문에 큰 DB에서 빠른 검색을 위해 차원 축소 필수적임.
- 256 Node 의 FC layer 를 이용하여, sigmod activation 후 feature extract 함
  threshold value 0.5로 설정, 0.5 이상이면 1 아니면 0으로 입력
- Hamming distance measure로 주어진 쿼리 wafer map에 적용

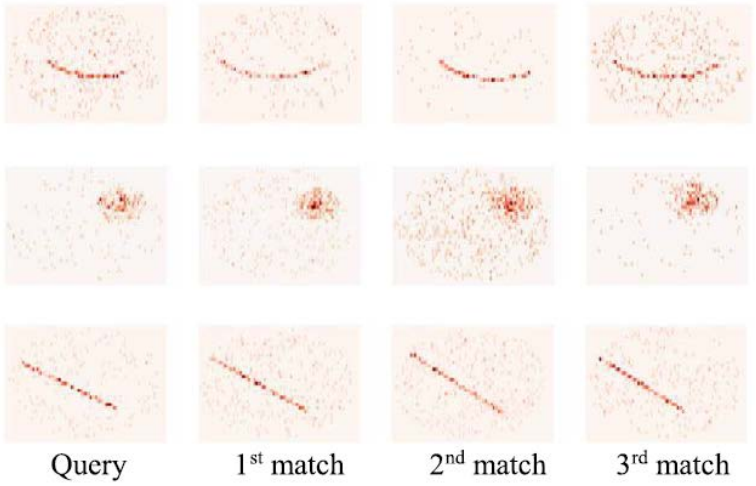


Defect pattern 을 이용한 Wafer 분류 및 유사 Wafer 검색

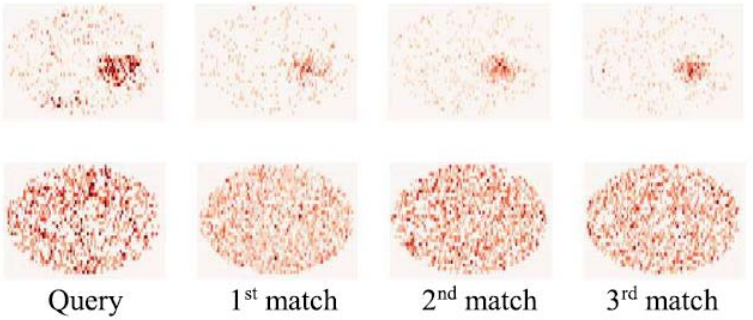
Wafer Map Image Retrieval

- defect 클래스 별 wafer map image 검색 결과
- 각 row별 첫번째 이미지는 쿼리 이미지, 나머지는 top 3 연관 검색 이미지

6,600 Simulated wafer map, Error rate 0.36%



1,191 Real wafer map, Error rate 3.7%



Defect pattern 을 이용한 Wafer 분류 및 유사 Wafer 검색

Conclusions

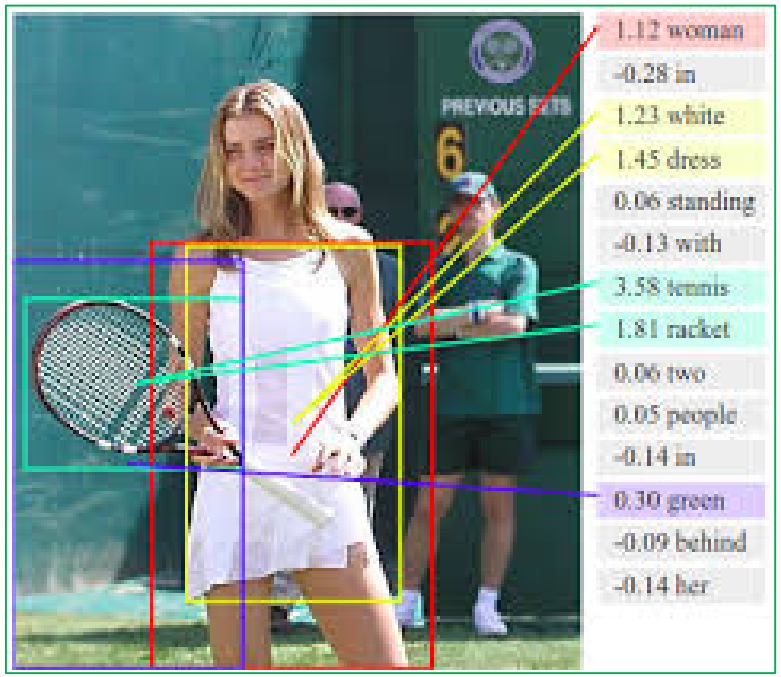
- CNN을 이용한 웨이퍼 맵 패턴 분류 및 웨이퍼 맵 이미지 검색 방법을 제시함.
- Simulated Wafer map 을 이용하여 small data 문제가 있는 상황에서도 학습을 잘 시킴.
- CNN 모델의 FC 레이어에서 생성 된 바이너리 코드를 사용하여 CNN 기반 이미지 검색의 효율성과 성능 개선을 확인 함.

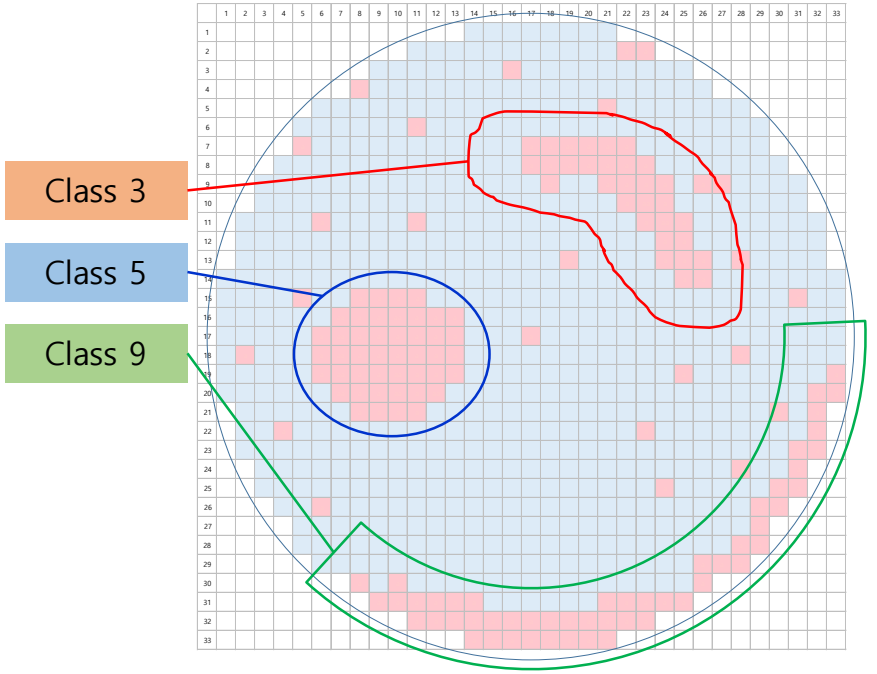
Limitations

- CNN model 학습을 위하여, 사용자가 생성한 22 class 의 simulated wafer map 을 사용.
  실제 상황을 더 잘 반영하기 위하여, 실제 data 를 이용하여 학습하는 것이 더 타당함.
  Data augmentation 또는 GAN(generative adversarial networks) 을 이용하여 학습력을 높일 수 있음.
- 본 CNN model 은 single label classification model 로써, Wafer 내에 다양하게 발생하는 문제에 대하여 대응이 어려움. Wafer 내에 있는 군집 pattern 을 기준으로 각각 class 를 할당할 수 있는 multi label classification model 을 이용한다면 더욱 좋은 분석 방안을 제시 할 수 있음.

Defect pattern 을 이용한 Wafer 분류 및 유사 Wafer 검색

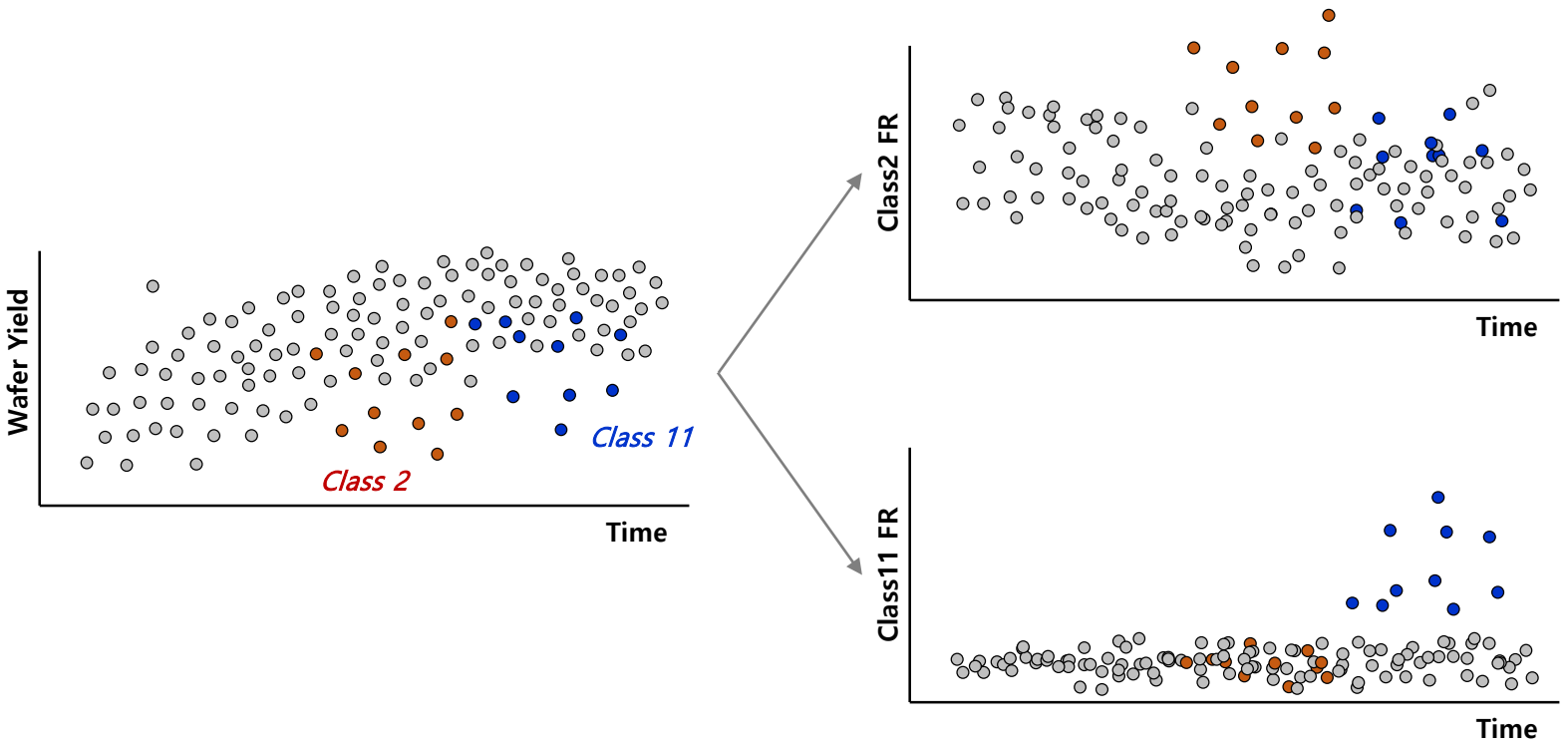
- **Multi label classification 적용 예**





Defect pattern 을 이용한 Wafer 분류 및 유사 Wafer 검색

- **Multi label classification 적용 예**



감사합니다.