

2022 대한안전경영과학회 추계학술대회

Grad-CAM 알고리즘을 통한 반도체 웨이퍼 불량 분석 및 원인 분석

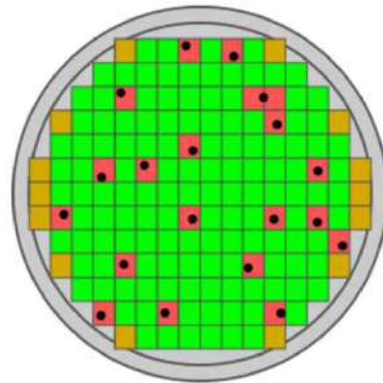
홍용민¹, 이현정², 김유림¹, 손준영¹, 한상혁¹, 성시열¹, 김예린³, 강성우¹
¹인하대학교 산업경영공학과, ²인하대학교 신소재공학과, ³인하대학교 경영학과

Contents

1. 서론	1.1 연구 동기 및 목적
2. 이론적 배경	2.1 Grad-CAM
3. 실험 방법 및 결과	3.1 데이터 수집 3.2 데이터 전처리 3.3 모델학습 3.4 Grad-CAM 3.5 Denoising
4. 결론	4.1 결론 및 향후 연구

서론 | 1.1 연구 동기 및 목적

반도체 EDS 공정: 전기적 특성검사를 통해 개별 Chip들이 원하는 품질 수준에 도달했는지 확인하는 공정
→ 반도체의 미세화와 수요 증가로 인해 엔지니어의 판단으로 수행되는 방법의 한계에 도달



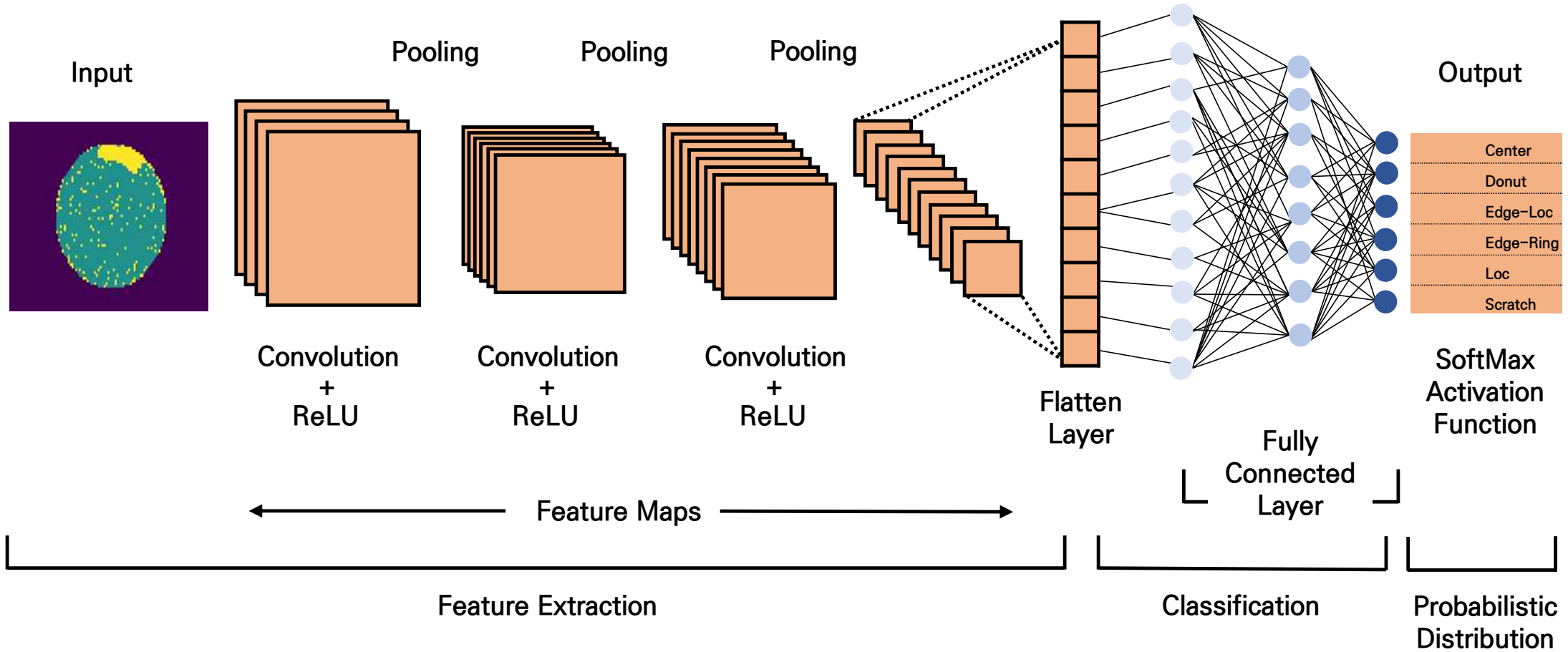
[Wafer Bin Map]

- : 양품의 Chip
- : 불량품의 Chip

AI를 활용하여 웨이퍼 맵 불량 패턴을 분류하고, Grad-CAM 알고리즘으로 주요 불량을 파악하고자 한다.

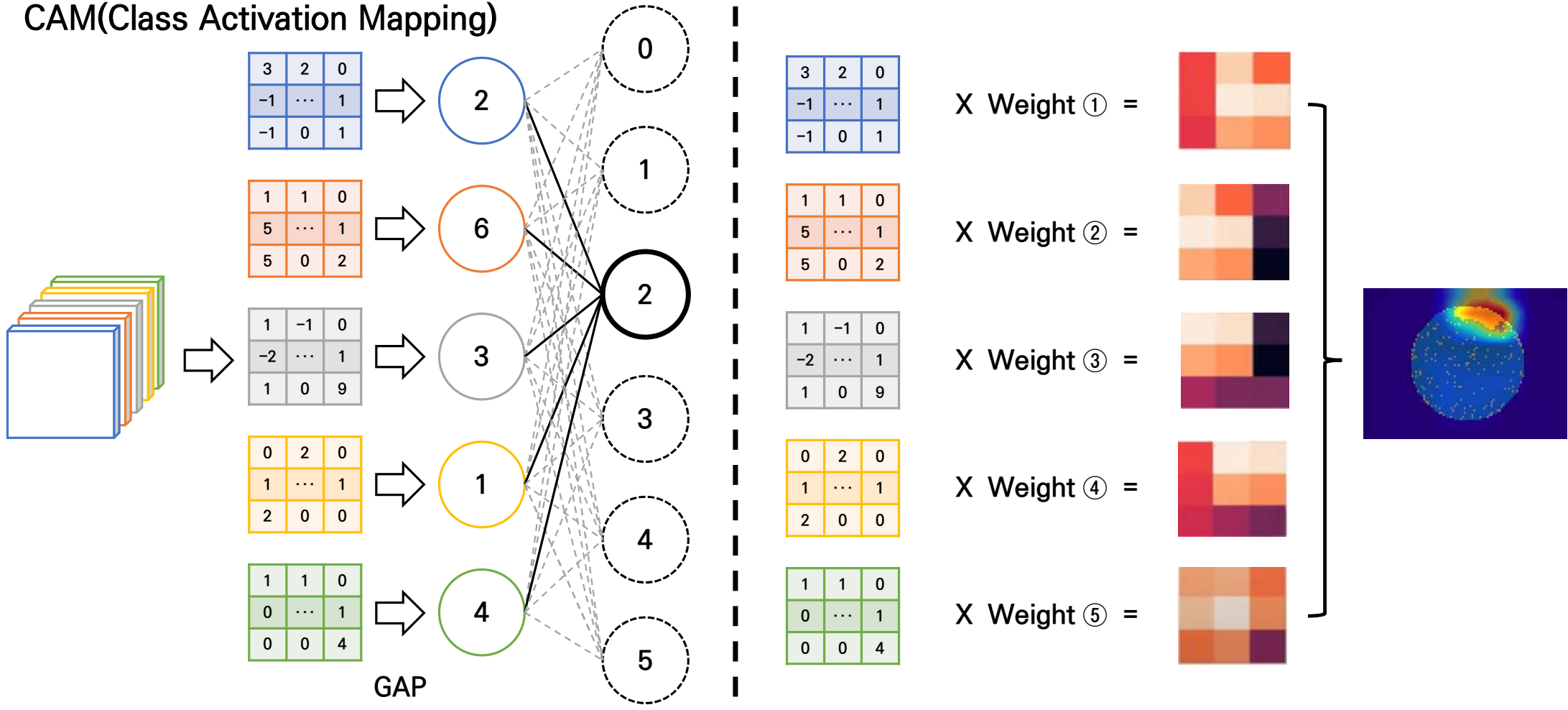
이론적 배경 | 2.1 Grad-CAM

CNN(Convolution Neural Network)



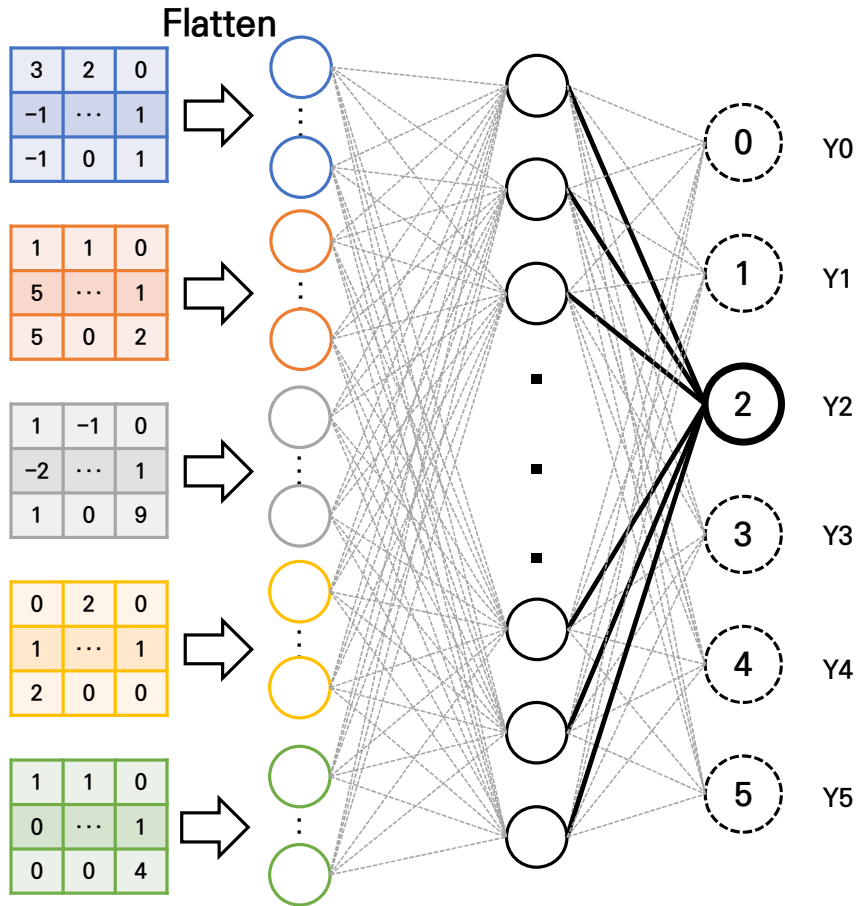
이론적 배경 | 2.1 Grad-CAM

CAM(Class Activation Mapping)



이론적 배경 | 2.1 Grad-CAM

Grad-CAM(Class Activation Mapping)



3	2	0
-1	...	1
-1	0	1

 $\times \frac{1}{z} \sum_i \sum_j \frac{\partial y^c}{\partial A_{i,j}^1} =$

1	1	0
5	...	1
5	0	2

 $\times \frac{1}{z} \sum_i \sum_j \frac{\partial y^c}{\partial A_{i,j}^2} =$

1	-1	0
-2	...	1
1	0	9

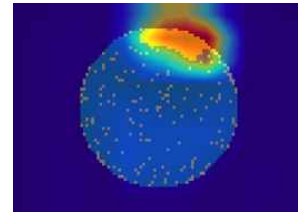
 $\times \frac{1}{z} \sum_i \sum_j \frac{\partial y^c}{\partial A_{i,j}^3} =$

0	2	0
1	...	1
2	0	0

 $\times \frac{1}{z} \sum_i \sum_j \frac{\partial y^c}{\partial A_{i,j}^4} =$

1	1	0
0	...	1
0	0	4

 $\times \frac{1}{z} \sum_i \sum_j \frac{\partial y^c}{\partial A_{i,j}^5} =$



실험 방법 및 결과 | 3.1 데이터 수집

- Kaggle에서 제공한 실제 제조 현장 46,393개의 lots에서 수집된 811,457 웨이퍼 맵을 포함한 데이터이다.
- WaferMap, DieSize, LotName, WaferIndex, TrainTestLabel, FailureType column으로 구성되어 있다.
- FailureType의 종류는 8가지로, Center, Donut, Edge-Loc, Edge-Ring, Loc, Random, Scratch, Near-full이 이에 해당한다.

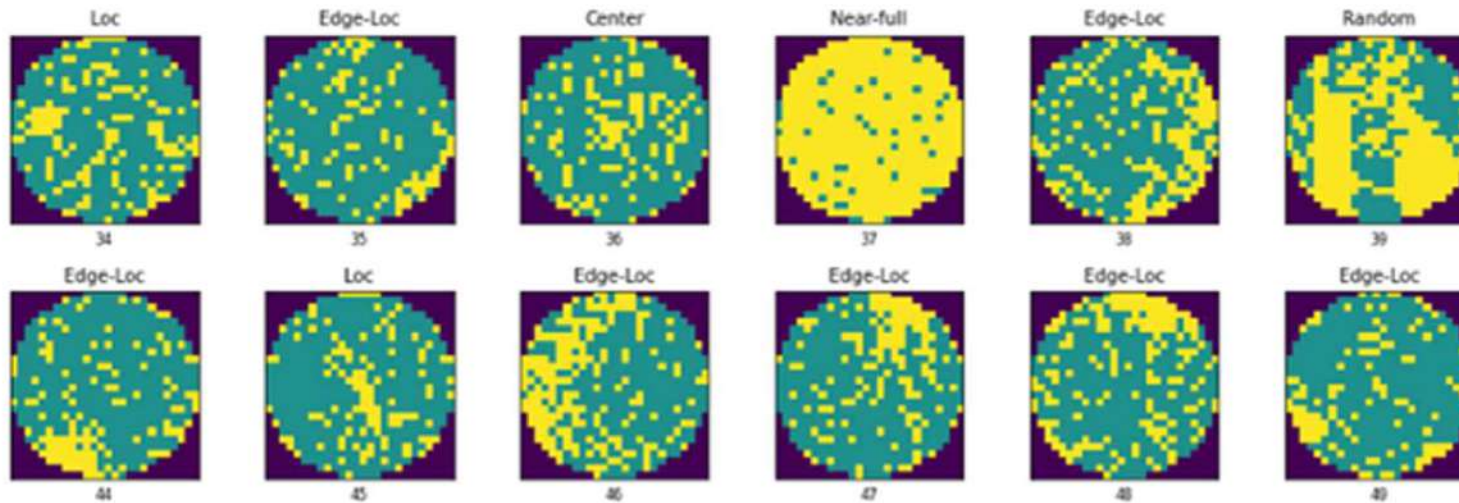


Figure 1. Wafer Image Data

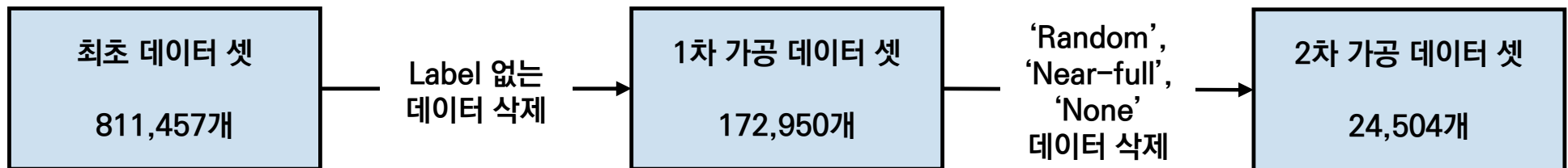
실험 방법 및 결과 | 3.2 데이터 전처리

1) Labeling 유무 확인

- Label이 없는 데이터 삭제

2) 유의미한 Failure Type 선별

- Random: 불량 발생이 Random하게 발생하여 경향성을 보이지 않는다.
- Near-full: 웨이퍼 전체에 불량이 발생한다.
- None: 불량이 발생하지 않는다.



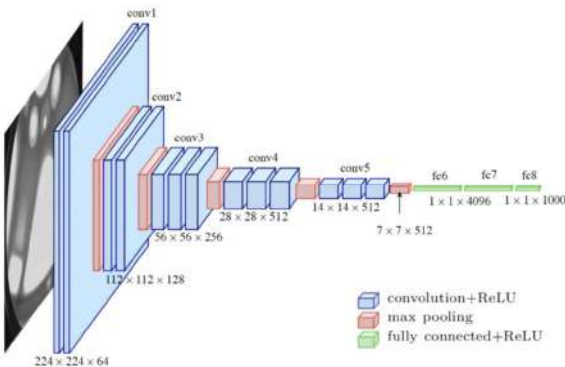
실험 방법 및 결과 | 3.3 모델 학습

1. VGGNet

- Network의 깊이를 깊게 만드는 것이 성능에 어떠한 영향을 미치는지 확인하기 위해 설계된 모델
- 13개의 Convolution Layer와 3개의 Fully Connected Layer로 구성

장점

- 깊이를 늘렸을 때 결정 함수의 비선형성 증가, 학습 파라미터 수 감소
- Network 깊이가 깊어질수록 객체 인식 Error가 감소하여 성능 향상

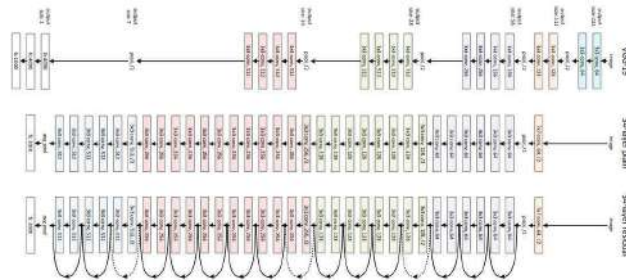


2. ResNet

- Network의 깊이가 깊어짐에 따라 발생하는 Overfitting, Vanishing Gradient 등의 문제를 해결하기 위해 Residual Learning을 도입한 모델
- Weight Layer 2개를 거치기 이전의 입력인 x 를 더해주어 $H(x) = F(x) + x$ 로 결과를 도출

장점

- 기존의 복잡한 문제 해결 가능
- Layer의 수가 증가하여도 정확도를 쉽게 얻을 수 있음

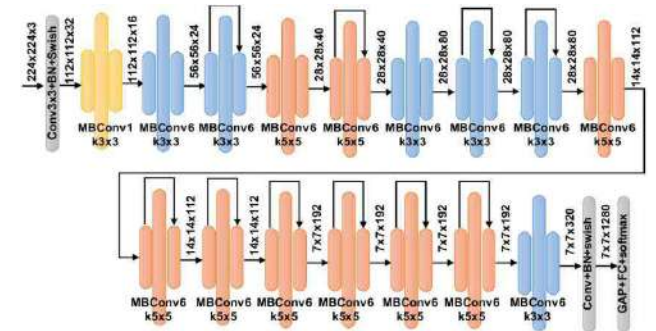


3. EfficientNet

- Image Classification에 대해, 기존보다 훨씬 적은 파라미터 수로 더욱 좋은 성능을 나타낸 모델

장점

- Network의 Depth를 깊게 만드는 방법, Channel Width를 늘리는 방법, Input Image의 해상도를 올리는 방법을 효율적으로 조절할 수 있는 Compound Scaling 방법을 제안함으로써 좋은 성능을 보임



실험 방법 및 결과 | 3.3 모델 학습

- CNN(Convolutional Neural Network)은 이미지를 분류하기 위해 개발된 알고리즘으로, 데이터에서 이미지를 직접 학습하고 분류한다.
- CNN 기반 모델인 VGGNet, ResNet, EfficientNet 모델을 사용하여 데이터를 학습하고, 불량 패턴을 분류한다.
- Epoch 수는 총 5회 수행하였으며 모델별 정확도는 다음과 같다.

	Epoch 1	Epoch 2	Epoch 3	Epoch 4	Epoch 5
VGGNet Accuracy	42.43	79.94	75.47	75.47	87.82
ResNet Accuracy	81.82	65.66	90.88	92.42	95.28
EfficientNet Accuracy	85.08	93.77	94.12	95.21	94.69

실험 방법 및 결과 | 3.4 Grad-CAM

- Grad-CAM은 설명가능한 인공지능인 XAI 중 하나로, AI의 Image Classification을 수행하는 데에 있어 높은 가중치를 주어 분류한 픽셀들을 선정하는 과정이다.
- 웨이퍼 맵 불량 패턴 Classification 과정에서 근거가 되는 픽셀들을 Grad-CAM을 통해 확인한다.
- 특정 불량이 빈번하게 나타나는 위치들을 파악한다.

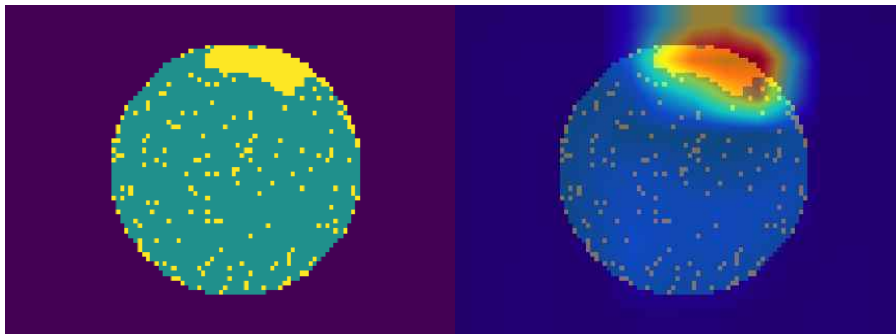


Figure 2. Original Data & Edge-Location Grad-CAM

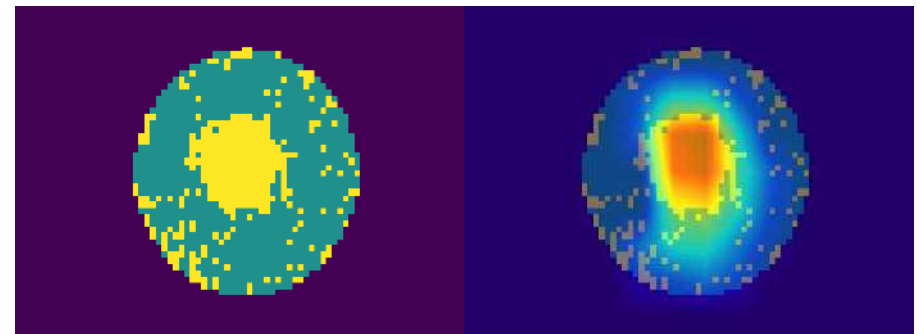


Figure 3. Original Data & Center Grad-CAM

기존 Image Denoising

각 이미지 내의 패턴을 바탕으로 Noise를 제거

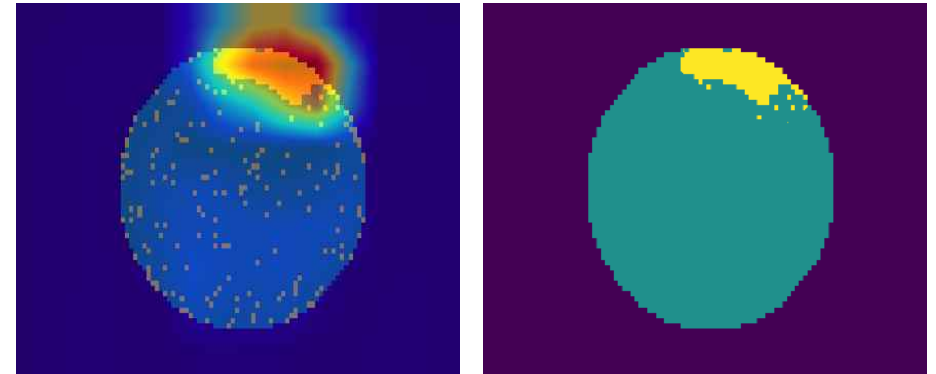
=> Median Blur 방법을 통한 Denoising 수행



개선 Image Denoising

Grad-CAM 알고리즘을 통한 Denoising

=> Dataset 전체를 고려한 Denoising 방법



결론

3가지 모델을 이용한
데이터 학습



VGGNet



ResNet



EfficientNet

Grad-CAM 알고리즘을 통한
불량 원인 분석

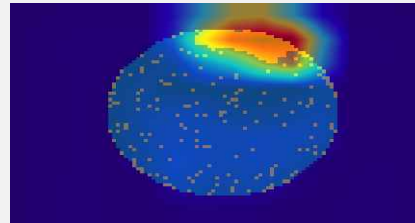
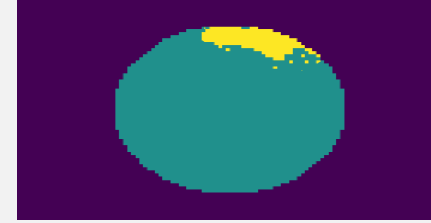


Image Dataset을 고려한
Denoising 수행



향후 연구

- Label이 없어 폐기한 약 60만 장의 데이터 활용을 위한 Auto Labelling 기법 개발
- Denoising 기법 Metric 개발을 통한 유효성 검증