
한정된 라벨 데이터를 극복하는 딥러닝 웨이퍼 결함 탐지: FixMatch 와 Autoencoder 준지도 학습 기법

박성진

University of Wisconsin-Madison

Computer Science & Data Science

(seongjinpark99@gmail.com)

[https://github.com/Seongjin74/Wafer](https://github.com/Seongjin74/Wafer-Defect-Classification)

-Defect-Classification

1. 서론

최근 딥러닝 기술이 급속도로 발전하면서, 전통적으로 사람이 수행해왔던 분류 및 예측 작업들이 자동화되고 있다. 특히, 이미지 분류, 결함 탐지, 예측 모델링 등의 분야에서 딥러닝은 뛰어난 성능을 보이며 기존 방법론을 대체하고 있다. 그러나 고성능 딥러닝 모델을 구축하기 위해서는 한정된 데이터를 효율적으로 활용하는 것이 매우 중요한 과제가 된다.

전통적인 딥러닝 학습 방식은 주로 지도학습(supervised learning)에 의존한다. 지도학습은 입력 데이터와 그에 상응하는 라벨이 모두 주어져야만 학습이 가능하며, 이로 인해 라벨링에 소요되는 인력과 비용이 큰 단점으로 작용한다. 예를 들어, 반도체 제조공정과 같은 복잡한 설비 환경에서는 대량의 라벨링된 데이터를 수집하는 데 많은 자원과 비용이 필요하다. 이와 같은 한계를 극복하기 위해, 최근에는 라벨링 비용과 인력 부담을 줄이면서도 우수한 성능을 내는 준지도학습(semi-supervised learning) 기법이 각광받고 있다.

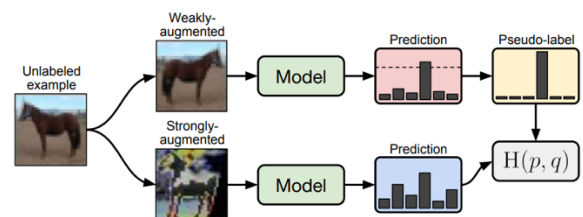
본 연구에서는 웨이퍼 결함 탐지를 위한 준지도학습 방법을 제안한다. 특히, 최신 준지도학습 기법 중 하나인 FixMatch 알고리즘을 활용하여 제한된 라벨링 데이터와 다량의 비라벨 데이터를 효과적으로 결합, 학습하는 방법을 제안한다. 아울러, autoencoder를 강한 증강(strong augmentation) 모듈로 활용하여, 원본 이미지의 재구성된 변형 이미지를 생성함으로써 모델이

pseudo-label과의 일관성을 유지하면서도 다양한 형태의 변형 데이터를 학습할 수 있도록 한다. 이러한 접근법은 단순히 autoencoder를 이용한 이미지 복원에 머무르지 않고, FixMatch의 강한 증강 단계에 통합되어 준지도학습의 효과를 극대화하는 데 기여할 것이다.

이 페이지에서는 먼저 준지도학습 및 autoencoder를 포함한 다양한 딥러닝 학습 기법의 이론적 배경과 기존 연구들을 검토하고, FixMatch 알고리즘과 autoencoder 기반 기법을 결합한 웨이퍼 결함 탐지 시스템을 상세히 설명한다. 이어서 제안한 방법의 성능을 다양한 실험을 통해 평가하고, 그 결과를 분석함으로써 향후 반도체 제조공정의 결함 탐지 및 예측 분야에 기여할 수 있는 가능성을 모색하고자 한다.

2. 관련 연구

최근에는 FixMatch 알고리즘이 준지도 학습 분야에서 단순하면서도 우수한 성능을 보이는 방법으로 각광받고 있다.

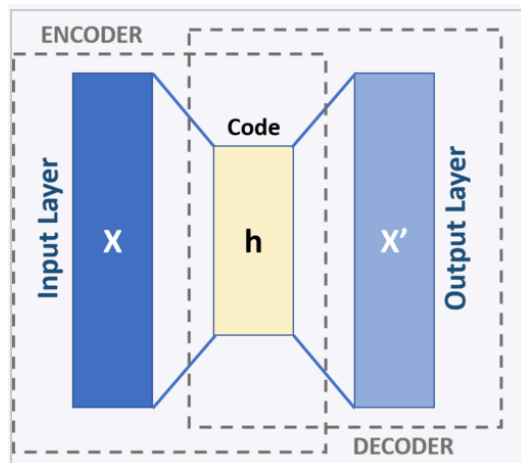


<그림 1> Diagram of FixMatch

Sohn 등(2020)은 FixMatch를 통해 약한 데이터 증강(weak augmentation)으로부터 생성된 예측

결과를 기반으로 pseudo-label을 생성하고, 이 pseudo-label이 일정 신뢰도(threshold, 예를 들어 0.90 이상)를 만족하는 경우에만 강한 데이터 증강(strong augmentation)된 이미지에 적용하여 모델을 학습시키는 방식을 제안하였다[1]. 구체적으로, 먼저 원본 이미지에 약한 증강을 적용하여 모델이 예측한 분포를 얻은 후, 가장 높은 확률을 가진 클래스를 pseudo-label로 선정한다. 이후, 동일 이미지에 강한 증강을 적용한 결과에 대해 pseudo-label과의 일관성을 유지하도록 모델 파라미터를 업데이트함으로써, 라벨이 부족한 상황에서도 효과적인 학습이 가능하도록 한다.

웨이퍼 결함 탐지 분야 역시 딥러닝 기술의 발전과 함께 전통적인 이미지 처리 기법을 넘어, 심층신경망(CNN) 기반의 방법들이 활발히 연구되고 있다. Wang 등(2018)은 CNN을 활용한 웨이퍼 결함 검사 시스템을 제안하여 기존 통계적 기법 대비 높은 탐지율을 달성하였으며, Kim 등(2019)은 준지도 학습 기법을 도입하여 제한된 라벨 데이터에서도 우수한 성능을 보이는 웨이퍼 결함 탐지 모델을 개발하였다[2][3]. 이와 함께, autoencoder를 활용한 연구들도 진행되고 있다.



<그림 2> A schema of an Autoencoder

Autoencoder는 입력 이미지를 압축하여 잠재 공간(latent space)으로 인코딩한 후, 이를 다시 원래의 이미지로 복원하는 디코딩 과정을 통해 이미지의 주요 특징을 효과적으로 추출한다. 이러한 방식은 노이즈 제거와 특징 재구성을 동시에 수행할 수 있어, 데이터 증강 및 결함 탐지의 강건성을 확보하는 데 기여한다. Choi 등(2021)은 autoencoder 기반의 접근법을 통해 반도체 제조공정에서 발생하는 미세한 결함까지도

효과적으로 검출할 수 있음을 입증하였다[4]. 이처럼, 준지도 학습 및 autoencoder 기반 기법은 웨이퍼 결함 탐지 분야에서 라벨 데이터의 한계를 극복하고 모델의 일반화 성능을 향상시키는 데 기여하고 있으며, 본 연구에서는 FixMatch 알고리즘과 autoencoder를 결합한 접근법을 제안하여, 기존 연구들을 포괄하는 성능 및 실용성을 확인하고자 한다.

3. 방법 제안론

본 연구에서는 웨이퍼 결함 탐지 문제를 해결하기 위해 준지도 학습 기반의 FixMatch 알고리즘과 autoencoder 기반 데이터 증강 기법을 결합한 새로운 방법론을 제안한다. 제안하는 방법은 데이터 전처리, 데이터 증강, 모델 설계 및 학습 과정으로 구성되며, 특히 약한 증강과 강한 증강의 조합, 그리고 이들을 FixMatch 프레임워크에 통합하는 데 중점을 두고 있다. 우선, 웨이퍼 맵 데이터는 process_wafer_map 함수를 통해 26×26 크기의 정규화된 3채널 이미지로 변환된다. 이 전처리 과정에서는 2차원 웨이퍼 맵의 각 픽셀 값을 미리 정의된 컬러맵에 따라 매핑하여, 원본 데이터의 결함 패턴을 효과적으로 부각시킨다.

데이터 증강 측면에서 제안하는 방법은 두 가지 증강 전략을 결합한다. 약한 증강은 원본 이미지의 기본 특성을 유지하도록 이미지 크기 조정과 소규모 회전 변환을 적용하며, 강한 증강은 보다 과감한 기하학적 변형을 통해 다양한 형태의 변형 데이터를 생성한다. 특히, 강한 증강에서는 autoencoder 기반 증강 기법이 핵심 역할을 한다. 학습된 autoencoder는 입력 이미지의 잠재 표현을 추출한 후, latent 공간에 노이즈를 추가하여 디코더를 통해 재구성된 이미지를 생성한다. 이 재구성 과정은 autoencoder가 웨이퍼 이미지의 내재된 특징을 학습하도록 하며, 재구성된 결과물은 후처리 과정을 거쳐 이산적인 값으로 변환된다. 또한, 강한 증강 단계에서는 autoencoder 증강 결과에 랜덤 90도 단위의 회전을 추가하여, 더욱 다양한 형태의 변형 이미지를 생성함으로써 모델이 다양한 변형에 대해 견고하게 학습할 수 있도록 한다.

FixMatch 알고리즘은 먼저 약한 증강을 적용한 이미지로부터 모델의 예측을 기반으로 pseudo-

label을 생성한다. 예측 확률이 사전에 정의된 임계값(예: 0.95) 이상일 때만 pseudo-label을 신뢰할 만한 라벨로 간주하며, 이후 동일한 원본 이미지에 강한 증강을 적용한 결과와의 일관성을 유지하도록 모델을 학습시킨다. 이 과정은 라벨이 제한된 상황에서도 대량의 비라벨 데이터를 효과적으로 활용하여 모델의 일반화 성능을 향상시키는 데 기여한다.

모델 설계 측면에서는 단순하면서도 효율적인 CNN 아키텍처인 SimpleCNN을 채택하였다. 이 네트워크는 3개의 합성곱 계층과 2회의 맥스 풀링 계층을 통해 이미지의 저차원 특징을 추출한 후 전결합 계층을 통해 최종 클래스를 예측하도록 구성된다. SimpleCNN은 구조가 간결하여 계산 복잡도가 낮고, 준지도 학습 환경에서 빠르게 수렴하는 장점을 가지며, 지도학습 및 준지도 학습 두 가지 실험 모두에서 사용되었다.

전체 학습 과정은 먼저 각 데이터셋에 대해 약한 증강과 강한 증강을 적용하는 것으로 시작된다. 지도학습 실험에서는 제한된 라벨 데이터를 활용하여 전통적인 CrossEntropyLoss 기반으로 모델을 학습시킨 반면, 준지도 학습 실험에서는 FixMatch 알고리즘에 따라 라벨 데이터와 비라벨 데이터를 동시에 사용하여 학습을 진행한다. 비라벨 데이터의 경우, 약한 증강을 통해 생성된 pseudo-label과 강한 증강을 적용한 이미지 간의 일관성을 유지하도록 추가 손실을 계산하며, 이때 pseudo-label의 신뢰도를 고려하여 가중치를 부여한다. 제안하는 방법은 웨이퍼 결함 탐지 분야에서 데이터 전처리, 증강, 준지도 학습 및 autoencoder 기반 재구성을 통합한 종합적인 접근법으로, 제한된 라벨 데이터를 효과적으로 활용하면서도 높은 탐지율과 일반화 성능을 달성할 수 있음을 목표로 한다.

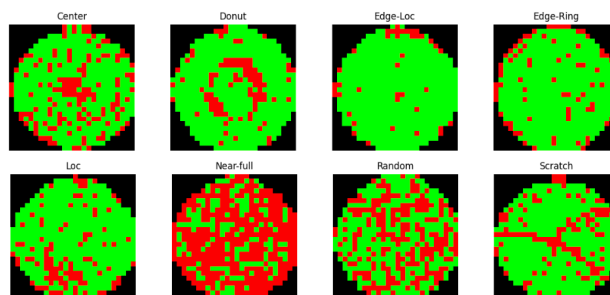
4. 실험

4.1 데이터셋

본 연구에서는 반도체 제조공정에서 수집된 WM-811K wafer map 데이터를 활용하여 제안한 방법론의 성능을 평가하였다. WM-811K 데이터셋은 Wu 등(2015)이 제시한 대규모 웨이퍼 맵 데이터셋으로, 실제 제조 환경에서 46,393 LOT에서 수집된 총 811,457개의 웨이퍼 맵을 포함하며, 결함

유형은 Center, Donut, Edge-Loc, Edge-Ring, Loc, Random, Scratch, Near-full, none 등으로 분류된다[5]. 이 데이터셋은 CC0 (Public Domain) 라이선스로 공개되어 있으며, 웨이퍼 결함 탐지 및 유사도 평가를 위한 연구에 널리 사용되고 있다. 전체 데이터셋은 우선 웨이퍼 맵 이미지와 라벨 정보를 전처리하여, 각 클래스별 균형 잡힌 데이터셋으로 분할하였다. 지도학습 실험에서는 각 클래스당 최대 100개의 학습 샘플과 50개의 테스트 샘플을 사용하였으며, 준지도 학습 실험에서는 동일한 100개의 라벨 샘플에 추가하여 라벨이 없는 데이터 10,000개를 활용하였다.

먼저, 데이터 전처리 과정에서 코드의 process_wafer_map 함수를 통해 입력된 웨이퍼 맵을 정규화된 3채널 이미지로 변환하였다. 이 함수는 웨이퍼 맵 내의 값(0, 1, 2)을 미리 정의된 컬러맵(0 → 검정, 1 → 초록, 2 → 빨강)으로 매핑한 후, 지정된 크기(26×26)로 리사이즈하여 [0, 1] 범위로 정규화된 이미지를 생성한다.



<그림3> 전처리후 웨이퍼 데이터 이미지

이후, Pandas 라이브러리를 활용하여 라벨이 있는 데이터를 선별하고, 각 클래스별로 80:20 비율로 학습/테스트 데이터셋을 구성하였다. 지도학습 실험의 경우, 각 클래스의 샘플 수를 제한함으로써 데이터 불균형 문제를 완화하고, 비교를 위한 기준을 마련하였다.

4.2 딥러닝 모델

모델은 단순 CNN 구조인 SimpleCNN을 사용하였으며, 세 개의 합성곱 계층과 두 번의 맥스 풀링을 통해 특징을 추출한 후, 전결합 계층을 통해 최종 클래스를 예측하도록 설계되었다. 학습은 Adam 옵티마이저(LR=1e-3)와 CrossEntropyLoss를 사용하였으며, 학습률은 10 에폭마다 0.5배씩 감소하는 스케줄러를 적용하여 총 50 에폭 동안 진행되었다.

4.3 지도학습

지도학습 실험에서는 Supervised DataLoader를

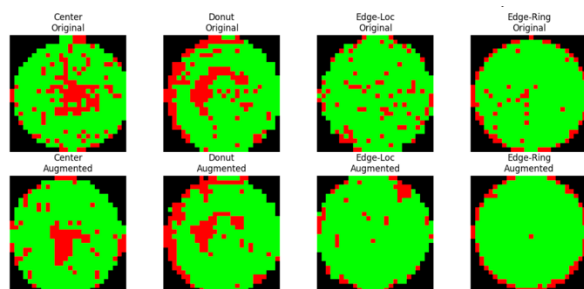
통해 모델을 학습시켰다. 학습 과정에서 에폭마다 손실 값이 점진적으로 감소하였으며, 최종 테스트 데이터셋에 대한 분류 정확도는 약 **73.72%**로 평가되었다. 또한, 클래스별 precision, recall, f1-score를 산출하여 모델의 세부 성능을 확인하였으며, 혼동 행렬을 통해 각 클래스 간의 분류 결과를 시각적으로 분석하였다.

Test Accuracy: 73.72%				
	precision	recall	f1-score	support
Center	0.88	0.84	0.86	50
Donut	0.90	0.92	0.91	50
Edge-Loc	0.70	0.64	0.67	50
Edge-Ring	0.87	0.90	0.88	50
Loc	0.59	0.48	0.53	50
Near-full	0.97	1.00	0.98	30
Random	0.85	0.92	0.88	50
Scratch	0.46	0.42	0.44	50
none	0.51	0.62	0.56	50
accuracy			0.74	430
macro avg	0.75	0.75	0.75	430
weighted avg	0.73	0.74	0.73	430

<그림4> 지도 학습 결과 표

4.4 준지도학습

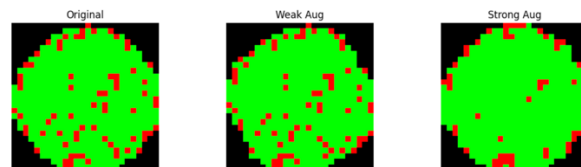
준지도 학습 실험에서는 먼저 약한 증강(weak augmentation)을 적용한 이미지로부터 모델의 예측 결과를 기반으로 pseudo-label을 생성하였다. 이때, 예측 확률이 사전에 정의된 임계값(예, 0.95) 이상인 경우에만 pseudo-label을 신뢰할 수 있는 라벨로 간주한다. 이후 동일한 원본 이미지에 대해서는 강한 증강(strong augmentation)을 적용하여, 이 강한 증강 결과가 pseudo-label과 일관되도록 모델을 학습시켰다. 강한 증강은 단순히 기하학적 변형만 적용하는 것이 아니라, autoencoder 기반 증강 모듈을 포함한다. 구체적으로, 학습된 autoencoder는 입력 이미지(26×26×3)를 인코더를 통해 13×13×64 크기의 잠재 표현(latent representation)으로 변환하고, 이 latent 공간에 소량의 노이즈(noise_std=0.05)를 추가한 후 디코더를 통해 재구성한다.



<그림 5> 클래스별 autoencoder 전후 웨이퍼 예시

재구성된 결과물은 후처리 과정을 거쳐 픽셀 단위의 이산적 변환을 적용받고, 추가적으로 90도

단위의 랜덤 회전이 적용된다. 이렇게 생성된 강한 증강 이미지는 원본 이미지의 핵심 결함 패턴을 유지하면서도 다양한 형태의 변형을 생성하여, FixMatch 알고리즘의 pseudo-label과의 일관성을 강제하는 역할을 한다.



<그림 6> 준지도 학습에 사용된 약한 증강과 강한증강 비라벨 웨이퍼 비교 예시

실험 초기에는 비라벨 데이터에서 pseudo-label의 활용 비율이 극히 낮았으나, 학습이 진행됨에 따라 에폭 50에선 전체 비라벨 샘플 중 약 16% 정도(예: 1002/6272)의 샘플에서 pseudo-label이 활용되었다. 준지도 학습 실험에서는 전체 테스트 정확도가 **76.05%**에 도달하였다.

[Semi-supervised Experiment Results]

Test Accuracy: 76.05%				
	precision	recall	f1-score	support
Center	0.90	0.90	0.90	50
Donut	0.90	0.92	0.91	50
Edge-Loc	0.76	0.62	0.68	50
Edge-Ring	0.96	0.88	0.92	50
Loc	0.65	0.48	0.55	50
Near-full	0.97	1.00	0.98	30
Random	0.81	0.92	0.86	50
Scratch	0.50	0.52	0.51	50
none	0.54	0.70	0.61	50
accuracy			0.76	430
macro avg	0.78	0.77	0.77	430
weighted avg	0.77	0.76	0.76	430

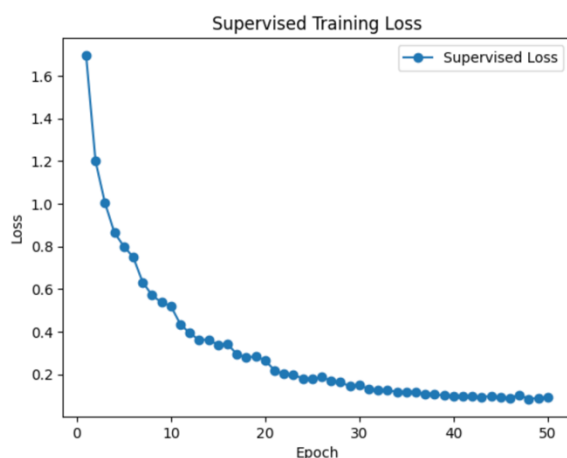
<그림7> 준지도 학습 결과 표

이와 같이, FixMatch 알고리즘은 제한된 라벨 데이터와 대량의 비라벨 데이터를 동시에 활용하면서 약한 증강을 통한 pseudo-label 생성과, autoencoder 기반 강한 증강(소량의 노이즈 주입 및 재구성, 추가 랜덤 회전 적용)을 통한 다양성 증대를 결합하여, 모델이 다양한 변형 데이터에 대해 견고하게 학습할 수 있도록 하였다. 또한, 제한한 autoencoder 기반 증강 기법은 원본 이미지의 내재된 특징을 효과적으로 보존하면서도, 추가적인 기하학적 변형을 통해 데이터의 다양성을 극대화하여, 준지도 학습 과정에서 모델의 일반화 성능 향상에 크게 기여한 것으로 분석된다.

5. 실험 평가

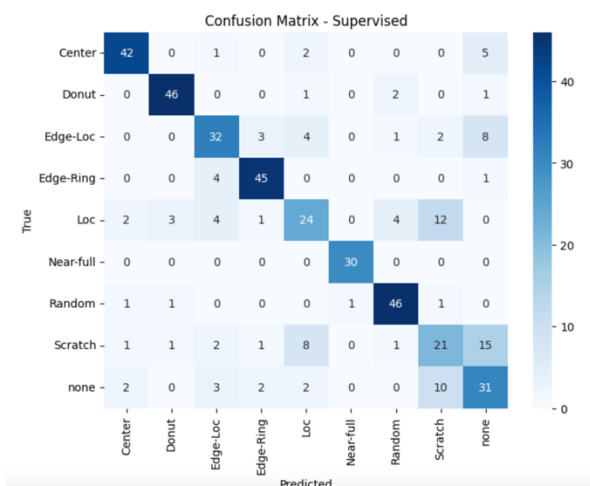
본 연구에서는 지도학습과 준지도 학습 두 가지 방법으로 학습한 모델의 성능 및 학습 과정을

수치적으로 비교·분석하였다. 지도학습 실험에서는 총 50 에폭 동안 SimpleCNN 모델을 학습하였으며, 초기 학습 손실은 약 1.73에서 시작하여 에폭 50에서 0.0920까지 감소하였다.



<그림 8> 지도 학습 손실 그래프

테스트 단계에서 지도학습 방식은 전체 테스트 정확도 **73.72%**를 달성하였다.

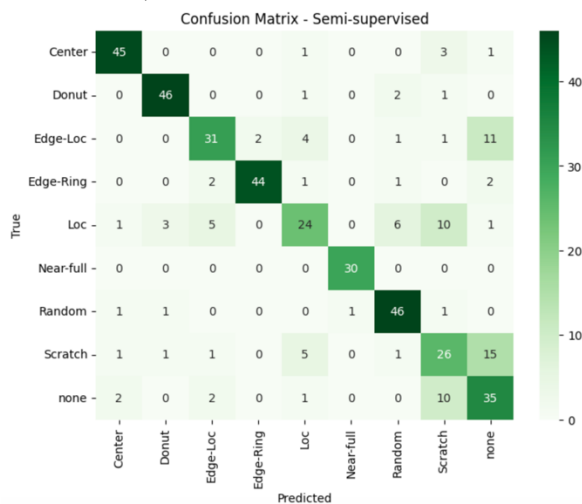


<그림 9> 지도 학습 Confusion Matrix

클래스별 세부 성능을 살펴보면, Center 클래스는 precision 0.88, recall 0.84, f1-score 0.86으로 나타났고, Donut은 0.90/0.92/0.91, Edge-Loc은 0.70/0.64/0.67, Edge-Ring은 0.87/0.90/0.88, Loc은 0.59/0.48/0.53, Near-full은 0.97/1.00/0.98, Random은 0.85/0.92/0.88, Scratch는 0.46/0.42/0.44, none은 0.51/0.62/0.56의 성능을 보였다. 이 결과는 지도학습 방식이 제한된 라벨 데이터를 기반으로 안정적인 학습을 수행하였으나, 일부 클래스(특히 Scratch, none 등)에서는 상대적으로 낮은 분류 성능을 보였음을 시사한다.

준지도 학습 실험에서는 동일한 900개의 라벨 데이터에 추가로 10,000개의 비라벨 데이터를

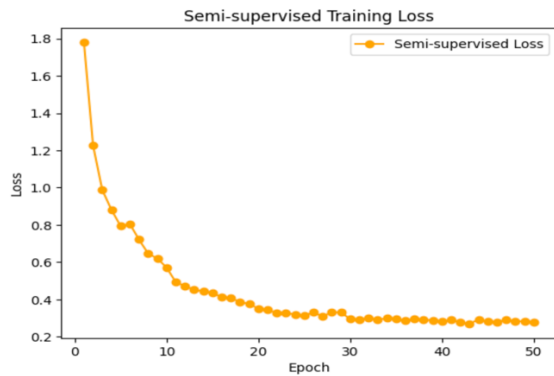
활용하여 FixMatch 알고리즘을 적용하였다. 준지도 학습의 경우, 초기 에폭에서는 비라벨 데이터에서 pseudo-label 활용 비율이 0.000에 가까웠으나, 학습이 진행되면서 에폭 50에 약 16% 정도의 활용 비율(예, 1002/6272)이 관측되었다. 최종적으로 준지도 학습 방식은 전체 테스트 정확도 **76.05%**를 달성하였으며, 클래스별 성능은



<그림 10> 준지도 학습 Confusion Matrix

Center(0.90/0.90/0.90), Donut(0.90/0.92/0.91), Edge-Loc(0.76/0.62/0.68), Edge-Ring(0.96/0.88/0.92), Loc(0.65/0.48/0.55), Near-full(0.97/1.00/0.98), Random(0.81/0.92/0.86), Scratch(0.50/0.52/0.51), none(0.54/0.70/0.61)로 나타났다.

지도학습과 준지도 학습을 직접 비교하면, 준지도 학습 방식은 전체 테스트 정확도를 약 2.3% 개선시켰으며, 특히 Center와 Edge-Ring 클래스에서 f1-score가 각각 0.86에서 0.90, 0.88에서 0.92로 약 4% 개선되는 등, 상대적으로 낮은 성능을 보이던 클래스(예, Scratch 클래스의 경우 0.44에서 0.51로 약 7% 개선이 있었고, none 클래스는 0.56에서 0.61로 5% 개선됨)에서 유의미한 향상을 가져왔다. 다만, Random 클래스는 precision이 다소 하락하였으나 전체 f1-score는 변화가 없었고, Edge-Loc의 경우 precision은 개선되었으나 recall에서 약간의 감소로 인해 f1-score 변화가 미미한 수준이었다.



<그림 11> 준지도 학습 손실 그래프

또한, 학습 손실 곡선 및 pseudo-label 활용 비율 모니터링 결과, 준지도 학습에서는 비라벨 데이터를 통한 추가 학습 효과가 점진적으로 나타나면서 모델의 수렴 속도와 일반화 성능을 향상시킨 것으로 분석된다. autoencoder 기반 강한 증강은 입력 이미지의 잠재 표현을 효과적으로 재구성하고, 추가적인 랜덤 회전을 통해 다양한 변형 이미지를 생성함으로써, FixMatch 알고리즘이 pseudo-label과의 일관성을 유지하도록 돕고, 결과적으로 제한된 라벨 환경에서도 비라벨 데이터의 이점을 극대화하는 역할을 수행하였다.

종합하면, 준지도 학습 방식은 지도학습에 비해 전체 정확도와 여러 클래스의 성능 지표에서 개선을 보여주었으며, 특히 autoencoder 기반 강한 증강과 FixMatch 알고리즘의 결합이 모델의 일반화 능력 향상에 크게 기여했음을 확인할 수 있었다. 이러한 결과는 본 연구에서 제안한 방법이 웨이퍼 결함 탐지 문제에 대해 제한된 라벨 데이터를 효과적으로 활용할 수 있는 실질적인 개선 효과를 제공함을 시사한다.

6. 결론 및 향후 연구

본 연구에서는 반도체 제조공정에서의 웨이퍼 결함 탐지를 위해, 제한된 라벨 데이터를 효과적으로 활용하는 준지도 학습 기법을 적용하였다. 제안한 방법은 FixMatch 알고리즘과 autoencoder 기반 데이터 증강 기법을 결합하여, 라벨이 부족한 상황에서도 모델의 일반화 성능을 향상시키는 것을 목표로 하였다. 실험 결과, 지도학습 실험에서 73.72%의 테스트 정확도를 달성한 반면, 준지도 학습 기법을 적용한 경우 76.05%의 테스트 정확도를 기록하여, 비라벨 데이터를 효과적으로 활용할 수 있음을 확인하였다.

또한, 클래스별 평가 지표와 혼동 행렬 분석 결과는, 특히 결함 패턴이 미세한 일부 클래스에서 준지도 학습의 효과가 더욱 두드러짐을 보여주었다.

향후 연구에서는 autoencoder 기반 강한 증강 기법과 FixMatch 준지도 학습의 결합 효과를 극대화하기 위해 여러 개선 방안을 도입해야 한다. 우선, 현재 latent 공간에 고정된 noise_std 값을 단순히 적용하여 노이즈를 추가하는 방식 대신, 하이퍼파라미터 튜닝을 통해 노이즈의 분포와 크기를 동적으로 조절하거나, 다양한 수준의 노이즈를 적용한 후 이를 ensemble 방식으로 결합하는 방법을 도입할 필요가 있다. 이를 통해 재구성 이미지의 다양성과 품질을 더욱 향상시킬 수 있다. 또한, autoencoder 모듈 자체를 보다 심층적이고 복잡한 네트워크 구조로 확장함으로써, 웨이퍼 이미지의 내재된 특징을 보다 정교하게 추출하고 재구성하는 방안을 마련해야 한다.

더불어, 준지도 학습에서 pseudo-label의 신뢰도 산출 방식을 개선하는 것도 중요한 과제로, 현재 사용 중인 고정 임계값(0.95)을 학습 초반과 후반의 모델 신뢰도 차이를 반영하여 동적으로 조절하거나, 예측 불확실성을 고려한 가중치 부여 방식을 도입하는 방법을 검토해야 한다. 이를 통해 비라벨 데이터가 모델 학습에 기여하는 효과를 더욱 극대화할 수 있다.

마지막으로, SimpleCNN과 같은 단순 모델도 유의미한 성능 개선을 보여주었으나, 향후에는 ResNet, DenseNet 등과 같이 보다 깊고 복잡한 네트워크 아키텍처를 도입하여 준지도 학습과 autoencoder 기반 강한 증강의 결합 효과를 극대화할 필요가 있다. 이와 같이 하이퍼파라미터 튜닝 및 다양한 개선 방안을 체계적으로 도입함으로써, 제한된 라벨 환경에서도 웨이퍼 결함 탐지의 정확도와 일반화 능력을 극대화할 수 있는 종합적인 프레임워크를 구축해야 한다.

참고문헌(Reference)

- [1] Sohn, K., Berthelot, D., Li, C.-L., Zhang, Z., Carlini, N., Cubuk, E.D., Kurakin, A., Zhang, H., & Raffel, C. (2020). FixMatch: Simplifying Semi-Supervised
- [2] Learning with Consistency and Confidence. Wang, Y., et al. (2018). Wafer Defect Inspection
- [3] Using Deep Convolutional Neural Networks.

Kim, S., et al. (2019). Semi-supervised Learning for
[4] Wafer Defect Detection. Choi, J., et al. (2021).
Autoencoder-Based Defect Detection in
Semiconductor Manufacturing.
Wu, Ming-Ju, Jyh-Shing R. Jang, and Jui-Long Chen.

[5] "Wafer Map Failure Pattern Recognition and
Similarity Ranking for Large-Scale Data Sets." IEEE
Transactions on Semiconductor Manufacturing 28,
no. 1 (February 2015): 1–12.