

# Vision Transformer와 유사도 기반 반도체 혼합 결함 유형 예측

1)최재준\*, 2)조성주\*

\*국립금오공과대학교 산업공학과

(39177) 경상북도 구미시 대학로 61

## Prediction of Mixed Defect Types in Semiconductors Using Vision Transformer and Similarity-Based Methods

Jae-June Choe\*, Seong-Ju Cho\*

\*Dept. of Industrial Engineering, Kumoh National Institute of Technology

---

1) First Author: Jae-June Choe, 2) Corresponding Author: Seong-ju Cho

\* Jae-June Choe(cjj1222@kumoh.ac.kr), Dept. of Industrial Engineering, Kumoh National Institute of Technology

\*\* Seong-Ju Cho(csj@kumoh.ac.kr), Dept. of Industrial Engineering, Kumoh National Institute of Technology

**요약:** 반도체의 집적회로(IC) 필요성과 발전으로 인하여 새로운 결함들이 생겨나고 있다. 반도체에서 결함은 단일 유형의 결함마다 원인이 담겨있다. 따라서 반도체 결함 탐지는 반도체 공정의 문제인식과 빠른 대처에 영향을 끼치는 중요 요소이자 제조업의 손실을 최소화할 수 있는 중요 정보이다. 특히, 그 중에서 혼합 유형의 결함은 기존처럼 1가지 결함 유형만을 판단하는 것이 아닌 2가지 이상의 결함 유형을 판단하는 것이기에 정확하고 빠르게 결함 유형을 판단하는 것이 기존 단일 유형의 결함보다 더욱 더 중요해지게 되었다. 따라서, 기존에 가지고 있는 단일 유형의 결함을 가지고 혼합 유형의 결함을 탐지 및 예측하는 것이 앞으로의 반도체 제조에서 중요하다고 판단했다. 해당 내용을 연구 주제로 잡고 이미지 분류 및 탐지에 있어서 CNN보다 대규모 데이터셋에서 효율적이라고 연구결과가 나온 Vision Transformer를 활용하기로 했다. 해당 모델을 구축한 뒤, 정상과 단일 유형의 결함 데이터를 학습하여 이중 혼합 유형의 결함 데이터에 대한 동일 차원의 임베딩 값을 추출하여 학습 데이터의 임베딩 값과 메트릭을 Euclidean 또는 Cosine 유사도 계산으로 활용하는 KNN을 통해 이중 혼합 유형의 결함 데이터의 결함 유형을 예측하는 Vision Transformer와 유사도 기반으로 결함 유형을 예측하는 프레임워크를 제안한다.

**ABSTRACT:** The advancement and necessity of integrated circuits (ICs) in semiconductors have introduced new types of defects. Each type of defect in semiconductors has its own underlying causes. Therefore, defect detection in semiconductors is crucial as it impacts problem recognition in semiconductor processes and enables quick responses, thus minimizing losses in manufacturing. Among these defects, mixed-type defects are particularly significant as they involve identifying more than one type of defect, making accurate and swift defect type determination more critical than for single-type defects. Consequently, detecting and predicting mixed-type defects using existing single-type defect data is deemed essential for future semiconductor manufacturing. As a research topic, we decided to use Vision Transformer, which has been found to be more efficient than CNN in handling large datasets for image classification and detection. After building the model, we will train it with normal and single-type defect data to extract embedding values of the same dimension for double mixed-type defect data. Using these embedding values, we propose a framework that predicts the defect type of double mixed-type defect data through KNN, utilizing metrics such as Euclidean or Cosine similarity to calculate similarities with the embedding values of the training data.

**Keywords :** *Mixed-Type Defects, Vision Transformer, Defect Detection, Semiconductor Manufacturing, K-Nearest Neighbors (KNN)*

## 1. 서론

### 1-1. 반도체 제조

웨이퍼(Wafer) 제조는 수많은 복잡한 과정을 거치는 반도체 제조 공정의 절차입니다. 웨이퍼 제조 후에는 각 웨이퍼의 다이(Die)는 전기 프로브(probe)로 테스트하고 웨이퍼 맵(wafer map)은 각 칩(chip)에 대한 테스트 결과를 기록합니다. 해당 과정을 반도체 공정에서 EDS(Electrical Die Sorting) 공정이라고 한다 [1]. 웨이퍼 맵에서 결함 패턴을 인식하면 결함의 원인을 쉽게 찾을 수 있기에 웨이퍼 제조 공정을 최적화하고 수율 향상에 있어서 도움이 된다[2]. 이러한 상황에서 최근 반도체 제조 공정은 집적회로(IC) 산업의 필요성과 발전에 따라 웨이퍼에 더 많은 회로를 집적하기 위하여 제조 공정이 점점 복잡하게 변화하고 있습니다. 이에 따라 자동으로 각각의 반도체 제조 공정이 점점 더 복잡해지고 있습니다. 이러한 상황 때문에 기존보다 반도체의 결함 빈도 수도 증가하고 있다. 여기서 발생하고 있는 또 다른 상황은 바로 웨이퍼에서 단일 유형의 결함이 발생하는 것이 아닌 2개 이상의 단일 유형의 결함이 혼합된 다중 유형의 결함이 발생하고 있다 [3]. 이러한 다중 유형의 결함의 가장 큰 문제점은 단일 유형의 결함과 달리 중복된 결함으로 인해 이에 대한 결함 패턴의 특성이 매우 복잡하다. 이는 단일 유형의 결함 탐지의 난이도와 달리 식별에 있어

더 어렵다. 이러한 상황에서 단일 유형의 결함 데이터에 대한 탐지 및 분류뿐만 아니라 다중 유형의 결함 데이터도 탐지 및 분류하는 것이 굉장히 중요해졌다. 단일 유형의 결함 탐지도 매우 중요하지만, 다중 유형의 결함 탐지도 매우 중요해졌다. 그 이유는 반도체의 결함은 종류에 따라 문제점이 있는 공정 과정을 보여주는 요소이기에 굉장히 중요하다. 결함 유형별 설명은 Fig. 1을 통해 확인할 수 있다. 이런 측면에서 다중 유형의 결함은 2가지 이상의 단일 유형의 결함이 존재하는 것이기 때문에 2가지의 결함을 정확하게 탐지하지 못한다면 반도체 제조와 공정의 문제점에 대한 정확한 파악과 해결책을 제시하기 어려워지기 때문이다. 따라서, 단일과 다중 유형의 결함을 모두 정확하게 탐지하고 분류하는 기술이 필요한 상황이다.

### 1-2. 선행연구

Byun, Yunseon and Baek, Jun-Geol (2020)은 사전 학습된 CNN과 초기 가중치 학습을 위하여 Convolutional Autoencoder를 사용했다. 이후 Softmax Function을 활용한 확률 기반 분류를 통해 혼합 유형의 결함 패턴 여부와 종류를 결정했다. 하지만, 단일 유형의 결함 분류의 정확도는 높았지만, 혼합 유형의 결함 분류 성능은 임계값 설정으로 인하여 상대적으로 낮았다 [5].

결함 패턴	텍스트 설명
Normal	일반 패턴은 웨이퍼에 결함이 없는 것이 특징입니다.
Center	센터결함(Central Defect)은 웨이퍼 중앙에 클러스터(cluster)가 생기는 것이 특징으로 플리즈마 영역이나 박막 증착, 산화 등의 문제로 인해 발생한다.
Dount	도넛 결함은 환형 클러스터가 특징이며 현상 공정 중 웨이퍼 표면에 남아 있는 용해된 포토레지스트 고형물이 재증착되어 발생합니다. 중앙 세척 단계에서 웨이퍼 중앙의 결함이 줄어들기 때문입니다.
Edge_loc	Edge loc은 웨이퍼 가장자리에 국부적인 클러스터가 나타나는 것을 특징으로 하며 확산 공정 중 가열이 고르지 않거나 식각 중 문제로 인해 발생합니다.
Edge_ring	에지 링 결함은 웨이퍼 주변에 링 모양의 클러스터가 나타나는 것을 특징으로 하며, 급속 열 어닐링 공정에서 에지의 비정상적인 온도 제어나 리소그래피 중 문제, 저장 노드 공정에서 충전 정렬 불량으로 인해 발생합니다.
Loc(Localized)	Loc 결함은 국부적인 클러스터가 특징이며 주어진 기계에서 과도한 진동으로 인해 발생하여 충분한 입자 또는 결정질 이질성이 발생합니다.
Near_Full	Nearfull Defect는 웨이퍼의 다이 중 90% 이상이 불량인 경우를 말합니다.
Scratch	스크래치 불량은 직선과 곡선이 특징이며, 배송 및 취급 과정에서 사람의 실수나 화학적 기계적 연마 과정에서 패드가 경화되어 발생합니다.
Random	무작위 결함은 패턴 중 하나에 기인할 수 없는 하나 이상의 결함이 특징이며, 제조 환경 요인에 의해 발생하며 체계화되지 않습니다.

Fig. 1. 단일 유형의 결함 데이터의 발생원인 [4]

Shim, J. and Kang, S. (2023)은 CNN을 사용하여 단일 웨이퍼 맵을 학습하고 혼합 결함 웨이퍼 맵을 분류했다. 해당 논문에서는 데이터 합성은 Mixup, Random Rotation, Noise Filtering을 통해 단일 유형의 결함 웨이퍼 맵을 합성하여 혼합 유형의 결함 웨이퍼 맵을 생성했다. 이를 통해 혼합 결함 웨이퍼 맵이 없는 경우에도 높은 예측 성능을 보였다 [6].

Yu, N., Chen, H., Xu, Q., and Hasan, M. M. (2021)은 적은 수의 학습 데이터를 사용하여 새로운 결함 패턴을 학습하고 분류하는 Few-Shot Learning을 사용했다. 이후 Hybrid Self-Attention Mechanism을 사용하여 결함 특징을 추출하고 메트릭 (Metric) 학습을 기반인 Prototypical Networks를 사용하여 새로운 결함 패턴을 인식하게 했다. 이를 통해 소수의 샘플을 사용하여 높은 분류 정확도를 달성했다 [7].

Wang, R., & Wang, S. (2023)은 웨이퍼 결함 패턴의 고장 진단을 위해 유클리드 거리를 사용한 유사한 결함 패턴을 검색하는 방법을 사용했다. 유사도 검색 방법을 통해 빠르고 효율적인 결함 진단을 제공했으며, 대규모 데이터셋에서도 높은 정확성을 보였다 [8].

Zhao, J., & Tang, L. (2023, June)은 웨이퍼 맵 결함 감지를 위한 BEiT(BERT Pre-Training of Image Transformers), FNet(Fourier Network), Vision Transformer(ViT) 및 Swin Transformer 등 네 가지 Transformer 기반 모델의 성능을 탐구했다. 이를 통해 Transformer 기반 모델이 웨이퍼 맵 결함 감지에 매우 효과적이라는 것을 보여줬다 [9].

선행연구를 통해서 단일과 다중 유형의 결함을 모두 정확하게 탐지하고 분류하는 연구를 확인했다. 하지만, 단일 유형의 결함 데이터만을 가지고 혼합 유형의 결함인 이미지 데이터의 결함 유형을 예측하는 주

제가 부족함을 확인하였다. 해당 주제는 Byun, Yunseon and Baek, Jun-Geol (2020)이 CNN과 Convolutional Autoencoder을 활용하여 진행했다.

본 연구에서는 Vaswani, A et al. (2017)와 Dosovitskiy et al. (2020)에서 밝힌 CNN보다 위치와 공간 정보를 유지하고 탐지할 수 있는 Vision Transformer를 사용하기로 했다 [10-11]. 이러한 이유는 웨이퍼의 결함 위치 및 공간 정보를 유지 및 탐지에 적합할 것이라는 생각이 들었기 때문이다. 이러한 ViT에 단일 유형의 결함 데이터를 학습시킨 뒤 유사도 기반으로 다중 유형의 결함 데이터의 결함 유형을 예측하는 연구 진행 및 프레임워크(framework)를 제안 및 실험하고자 한다.

### 1-3. 데이터셋

사용한 데이터 셋은 Kaggle이라는 데이터 분석 및 머신러닝에 대한 학습 플랫폼에 있는 Mixed-type Wafer Defect Datasets를 사용했다. 해당 데이터셋의 구조는 이미지 데이터를 의미하는 arr\_0과 해당 이미지의 결함 유형을 의미하는 arr\_1로 구성된 npz파일이다. 해당 파일에 존재하는 결함 유형은 Single-Types와 Two, Three, Four-Types Mixed가 존재한다. 하지만, 본 연구에서는 단일 유형의 결함 데이터를 가지고 혼합 유형의 데이터에 대한 결함 유형을 예측한 주제이기에 Single-Types와 Two-Types Mixed만을 사용하겠다. 사용하는 데이터셋의 원본은 Fig. 2를 통해 확인할 수 있다. 해당 데이터셋의 이미지 구조는 총 3가지의 픽셀값으로 구성되어 있다. Blank Spot을 의미하는 픽셀 0, 정상을 의미하는 픽셀 1, 결함을 의미하는 픽셀 2로 구성되어 있다. 해당 부분에 대한 설명은 Fig. 3에서 확인할 수 있다.

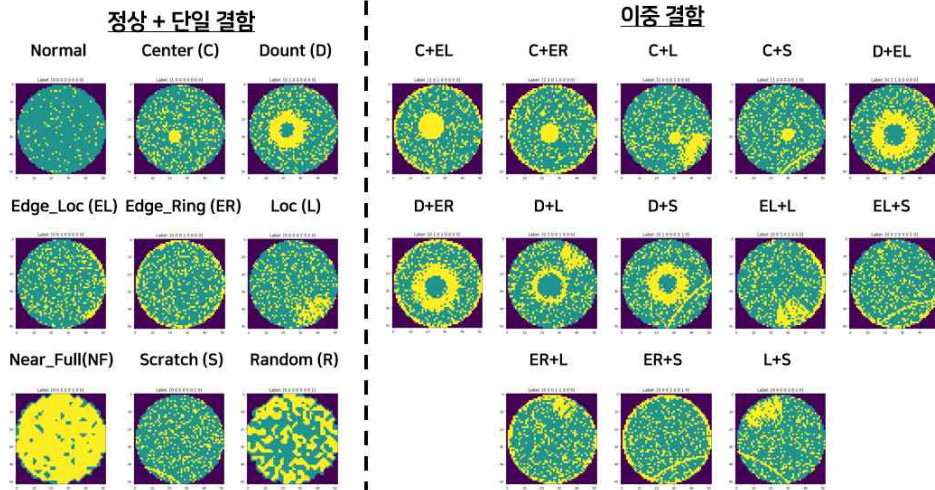
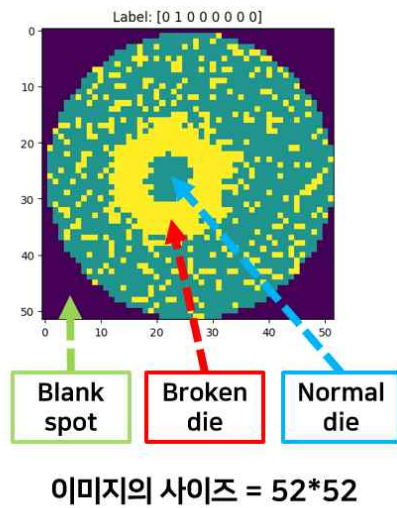


Fig. 2. 연구에서 사용할 단일과 이중 혼합 유형의 결함 이미지 데이터



정상(1가지) + 단일결함(8가지)

Label	유형	개수
Normal (정상)	[0 0 0 0 0 0 0]	1000
Center	[1 0 0 0 0 0 0]	1000
Donut	[0 1 0 0 0 0 0]	1000
Edge_Loc	[0 0 1 0 0 0 0]	1000
Edge_Ring	[0 0 0 1 0 0 0]	1000
Loc	[0 0 0 0 1 0 0]	1000
Near_Full	[0 0 0 0 1 0 0]	149
Scratch	[0 0 0 0 0 1 0]	1000
Random	[0 0 0 0 0 0 1]	866

이중결함(13가지)

Label	유형	개수
Center+Edge_Loc	[1 0 1 0 0 0 0]	1000
Center+Edge_Ring	[1 0 0 1 0 0 0]	1000
Center+Loc	[1 0 0 0 1 0 0]	1000
Center+Scratch	[1 0 0 0 0 1 0]	1000
Donut+Edge_Loc	[0 1 1 0 0 0 0]	1000
Donut+Edge_Ring	[0 1 0 1 0 0 0]	1000
Donut+Loc	[0 1 0 0 1 0 0]	1000
Donut+Scratch	[0 1 0 0 0 1 0]	1000
Edge_Loc+Loc	[0 0 1 0 1 0 0]	1000
Edge_Loc+Scratch	[0 0 1 0 0 1 0]	1000
Edge_Ring+Loc	[0 0 0 1 1 0 0]	1000
Edge_Ring+Scratch	[0 0 0 1 0 1 0]	1000
Loc+Scratch	[0 0 0 0 1 0 1]	1000

Fig. 3. 사용할 데이터의 이미지 구조와 개수

교하여 CNN과 ViT 모델의 성능을 비교하고자 한다.

## 2. 본론

### 2-1. 연구의 프레임워크

제안하는 연구의 프레임워크는 Fig. 4와 같다. 크게 데이터 수집, 데이터 전처리, ViT 모델 구축, 테스트 데이터에 대한 유형 예측으로 구성되어 있다. 사용할 유사도 계산 방식은 Euclidean과 Cosine으로 설정하였고, 이를 설계할 ViT 모델에 적용하여 각각 비교한 뒤, 마지막에는 앞서 선행연구에서 소개했던 Byun, Yunseon and Baek, Jun-Geol (2020)의 결과값과 비

### 2-2. 데이터 전처리

데이터 전처리는 Fig. 5의 과정으로 진행하였다. 먼저, 해당 데이터셋에 대한 설명이 적힌 문서에 없는 이미지 데이터의 픽셀값인 이상치에 대해 확인하였다. 그 결과 픽셀값이 3인 데이터가 존재하여 해당 이미지는 삭제하였다. 이후 데이터의 유형에 대한 개수 관련 불균형을 해소하기 위해서 데이터 증강기법을 통해 데이터의 각각의 유형을 1,000개에 도달하도록 적용하

였다. 증강 방법은 회전과 반전을 사용하였다. 회전은 90°, 180°, 270°를 적용했고, 반전은 좌우, 상하, 주대각선(이미지의 왼쪽 위에서 오른쪽 아래로 향하는 대각선), 부 대각선(이미지의 오른쪽 위에서 왼쪽 아래로 향하는 대각선)을 적용했다. 이를 통해 사용할 데이터의 각각의 유형은 1,000개 맞게 증강하였다. 해당 연구 주제에 맞게 정상과 단일 유형의 결함만을 모델의

학습 데이터로 사용하기에 총 9,000개의 데이터를 학습 데이터로 분류했고, 이중으로 된 13가지의 혼합 유형의 결함 데이터 총 13,000개의 데이터를 테스트 데이터로 분류했다. Fig. 6을 통해 확인할 수 있다.

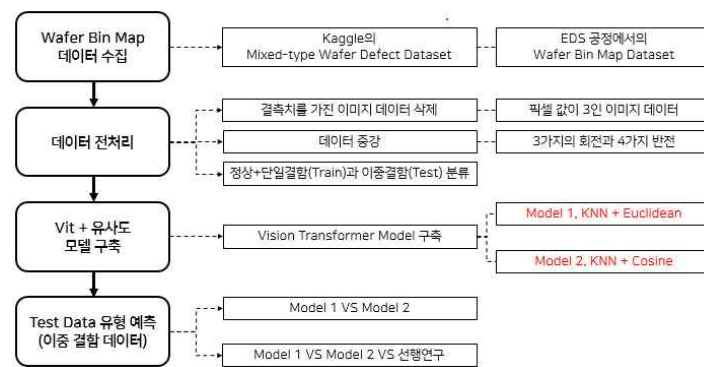


Fig. 4. 연구 프레임워크

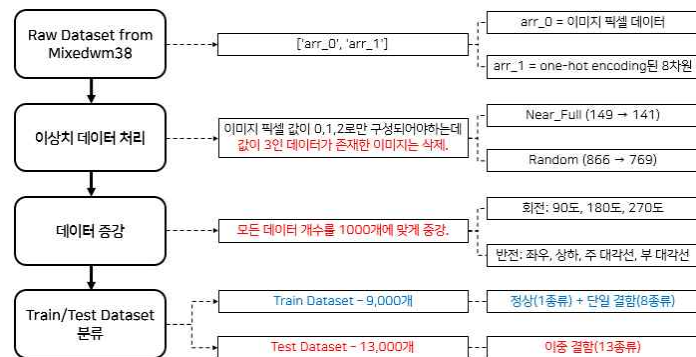


Fig. 5. 데이터 전처리 진행과정

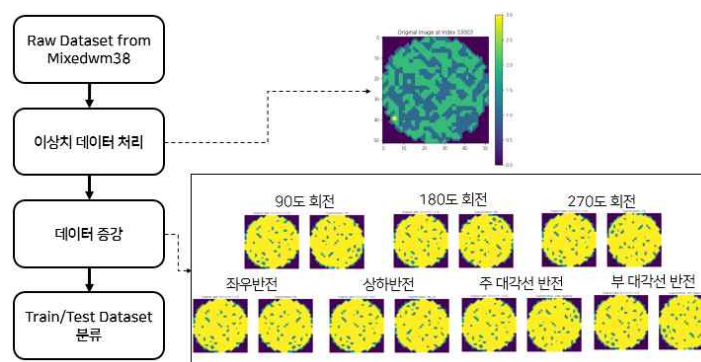


Fig. 6. 이상치 데이터 처리 및 증강 과정

## 2-3. ViT Model 구조 및 설계

ViT (Vision Transformer)는 Vanswani 등이 제안한 Transformer 모델(Model)의 원리를 확장한 모델 아키텍처(architecture)이다. 이는 자연어 처리 작업을 컴퓨터 비전(computer vision) 영역으로 확장하기 위해서 만든 아키텍처이다. Transformer의 주요 혁신은 반복 또는 컨볼루션 레이어(Convolutional Layer)가 필요하지 않고 모델이 입력과 출력 시퀀스(sequence) 간에 전역(global) 종속성을 도출할 수 있는 self-attention mechanism에 의존하는 것이다. 이러한 ViT 모델에서는 입력 이미지(image)는 고정된 크기의 패치(patch)로 나누어지며, 각 패치는 선형으로 임베드(embedding) 되고 위치 인코딩(encording)과 결합하여 공간 정보를 유지한다. 그런 다음 이러한 임베딩은 multi-head self-attention 및 feed-forward network 계층을 적용하는 Transformer에 의해 처리된다 [10-11].

ViT Model 구조설계는 이러한 내용을 참고하여 설계했다. 크게 Patch Embedding, Patch Encoder, Transformer Encoder, Training으로 구성되어 있다. [10-11]. Fig. 7을 통하여 설계한 Model 구조를 확인할 수 있다.

### 2-3-1. Patch Extraction

입력 이미지는 고정 크기의 패치(16×16)로 나누어진다. 이후 각 패치는 평탄화 (Flatten) 되어 재구성됩니다. 이후 size와 strides 파라미터를 통해 패치의 크기와 패치를 추출할 간격을 지정한다.

### 2-3-2. Patch Embedding

각 패치는 밀집 벡터 공간으로 Projection 되고 위치 인코딩으로 보강된다. 이 단계는 모델이 원본 이미지의 공간 정보를 유지하도록 한다. Projection 층은 각 패치를 고차원 공간으로 변환하고 Position\_embedding 층은 각 패치에 위치 정보를 추가한다.

### 2-3-3. Transformer Encoder Layers

모델은 임베드 된 패치를 처리하기 위해 여러 Transformer Encoder Layer를 적용한다. 각 레이어는 Layer Normalization을 통해 입력을 정규화하여 학습 안정성을 높인다. Multi-Head Self-Attention은 입력 시퀀스의 각 위치에 대해 다른 위치를 참조하여 정보를 추출한다. MLP 블록을 통해 각 위치에 독립적으로 적용되는 두 개의 밀집 레이어를 적용한다. Residual Connection은 입력과 출력을 더하여 그래디언트(gradient) 소실 문제를 완화한다. 최종적으로는 logits 함수에 Softmax Function을 적용하여 학습 데이터의 단일 결함의 유형 클래스에 따라 확률 분포를 생성한다. 이는 해당 클래스일 확률을 의미한다.

### 2-3-4. Multi-Layer Perceptron (MLP)

MLP Layer는 추가 학습 용량을 제공하고 모델 학습에 있어 반도체 결함의 복잡한 패턴을 학습할 수 있도록 한다. unit 수를 2048, 102로 구성하였다. Activation Function은 GELU를 사용했고, Dropout Layer를 포함한 밀집 레이어로 구성된다.

### 2-3-5. Training

해당 모델 학습은 지수 학습률 감소 스케줄을 설정한 Adam Optimizer를 사용했다. Checkpoint도 사용하여 Validation loss를 기준으로 최상의 모델을 저장하도록 했다. 초기 학습률은  $1 \times 10^{-4}$ 이고 1,000 step마다 0.9로 감소하게 스케줄을 생성했다. batch\_size는 128, epochs는 100으로 설정하였다. 해당 방법을 통해 학습이 진행됨에 따라 학습률을 감소시켜 안정적인 모델로 수렴하도록 한다.

### 2-3-6. Embedding Extraction

MLP에서 모델의 특정 레이어에서 96차원의 임베딩 벡터를 추출하여 Train Data와 Test Data의 임베딩 값에 대해 분석하도록 생성했다. 모델의 마지막 두 번째 레이어의 출력을 임베딩 벡터로 추출하도록 한다.

### 2-3-7. KNN 유사도 계산

ViT 모델을 통해 추출된 테스트 데이터의 96차원의 임베딩 벡터를 학습 데이터의 Embedding 값과의



유사도를 Euclidean과 Cosine으로 계산하고, 예측 결함 유형(레이블)을 생성할 수 있게 했다. KNN의 N은 5로 설정했으며 5개의 이웃 데이터에서 다수결로 2개의 서로 다른 결함 유형을 선택하도록 설정했다. 만약 5개에서 결정이 안 나면 N을 1씩 증가시켜서 결정하도록 했다. Fig. 8를 통해 이를 확인할 수 있다.

$$Euclidean\ Distance = \sqrt{\sum_{i=1}^d (a_i - b_i)^2} \quad (1)$$

$$Cosine\ Similarity = \frac{a \cdot b}{\|a\| \|b\|}$$

$$a \cdot b = \sum_{i=1}^d a_i b_i$$

$$\|a\| = \sqrt{\sum_{i=1}^d a_i^2}$$

$$\|b\| = \sqrt{\sum_{i=1}^d b_i^2}$$

(2)

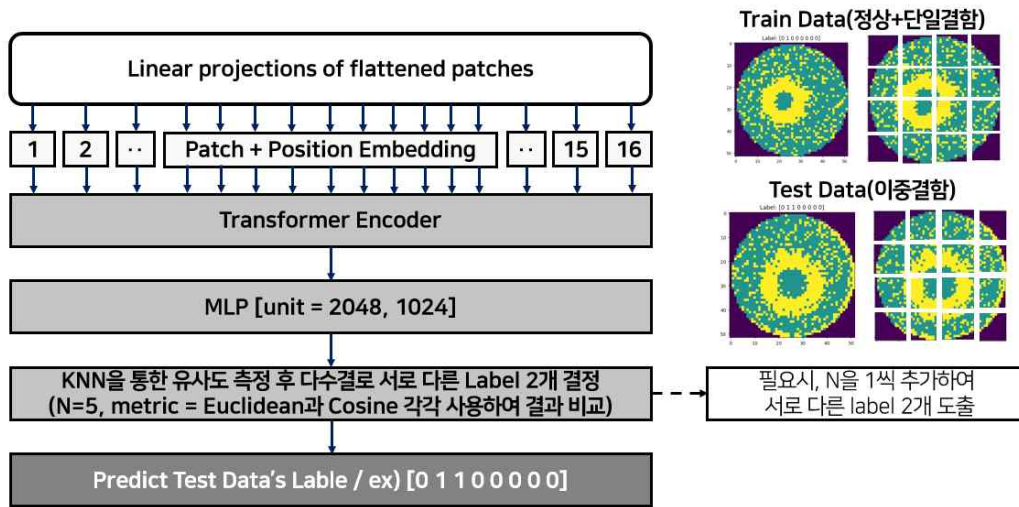


Fig. 7. ViT 모델 구조



**Algorithm 1** Calculate KNN Similarity

---

```

1: procedure CALCULATEKNNSIMILARITY(train_embeddings, test_embeddings, y_train, y_test, k = 5)
2:   results  $\leftarrow$  {}
3:   for metric in ['euclidean', 'cosine'] do
4:     nbrs  $\leftarrow$  NearestNeighbors(n_neighbors = k, algorithm = 'auto', metric = metric).fit(train_embeddings)
5:     distances, indices  $\leftarrow$  nbrs.kneighbors(test_embeddings)
6:     predictions  $\leftarrow$  []
7:     for i, test_embed in enumerate(test_embeddings) do
8:       knn_indices  $\leftarrow$  indices[i]
9:       knn_labels  $\leftarrow$  y_train[knn_indices]
10:      unique_labels  $\leftarrow$  np.unique(knn_labels, axis = 0)
11:      current_k  $\leftarrow$  k
12:      while len(unique_labels) < 2 do
13:        current_k  $\leftarrow$  current_k + 1
14:        nbrs.set_params(n_neighbors = current_k)
15:        distances, indices  $\leftarrow$  nbrs.kneighbors(test_embeddings)
16:        knn_indices  $\leftarrow$  indices[i]
17:        knn_labels  $\leftarrow$  y_train[knn_indices]
18:        unique_labels  $\leftarrow$  np.unique(knn_labels, axis = 0)
19:      end while
20:      label_counts  $\leftarrow$  np.sum(knn_labels, axis = 0)
21:      top_two_labels  $\leftarrow$  np.argsort(label_counts)[-2:]
22:      predicted_label  $\leftarrow$  np.zeros_like(knn_labels[0])
23:      predicted_label[top_two_labels]  $\leftarrow$  1
24:      predictions.append({'ActualLabel' : y_test[i], 'PredictedLabel' : predicted_label, 'KNN Labels' : unique_labels})
25:    end for
26:    results[metric]  $\leftarrow$  pd.DataFrame(predictions)
27:  end for
28:  return results
29: end procedure

```

---

**Fig. 8.** KNN을 이용한 Euclidean과 Cosine 유사도 계산

함 탐지 능력이 제대로 발휘되지 않음을 의미한다.

### 3. 결과 및 토론

#### 3-1. 연구결과

제안한 ViT 모델에 대해서 학습 데이터에 대해서 검증 데이터를 9:1로 분할하여 진행 결과는 Fig. 9에서 볼 수 있듯이 매우 좋게 나왔습니다. 하지만 이후 혼합 유형의 결함 데이터인 테스트 데이터에 대한 ViT와 Euclidean, Cosine 유사도는 Fig. 10처럼 좋지 못한 Accuracy, F1-Score, Recall, Precision을 보여 주었습니다. Fig. 11을 통해 ViT와 각 유사도에 대한 결함별 성능 지표를 확인할 수 있습니다. Fig. 11을 보면 Edge\_Ring, Near\_Full, Randomd에 대한 결함 탐지 능력은 긍정적인 수치를 보여줬습니다. 특히, Edge\_Ring은 Euclidean에서 상대적으로 더 좋은 수치를 보여줬습니다. 하지만 다른 유형의 결함에서는 굉장히 취약한 모습을 보여주었습니다. 특히, Edge\_Loc가 가장 낮은 탐지 능력을 보여주고 있습니다. Loc와 Scratch는 재현율(Recall)이 정밀도(Precision)에 비해 상대적으로 매우 낮았다. 이는 결

$$Accuracy = \frac{TP + TN}{TP + FN + FP + TN} \quad (3)$$

$$Precision = \frac{TP}{TP + FP} \quad (4)$$

$$Recall = \frac{TP}{TP + FN} \quad (5)$$

$$F1\ Score = \frac{Precision \times Recall}{Precision + Recall} \times 2 \quad (6)$$

단일 유형의 결함 데이터만을 모델에 학습시키고 혼합 유형의 결함 데이터를 예측하는 같은 주제를 진행한 Byun, Yunseon and Baek, Jun-Geol (2020)와 결과값을 비교해 보았다. 비교한 결과는 Fig. 12를 통해 확인할 수 있다. Fig. 12에서 Center+Edge\_Ring, Center+Scratch, Edge\_Ring+Scratch에서는 제안하는 ViT와 유사도 기반의 모델이 선행 연구보다 나은 수치를 가지고 있다. 하지만 Center+Edge\_Loc, Edge\_Loc+Scratch와 같이 Edge\_Loc와 혼합된 결함 유형에서는 선행연구보다 높은 탐지 및 예측률을 가지고 있음을 알 수 있다. 하지만 해당 두 유형의 수

치 차이보다 타 유형의 수치 차이가 더 크기 때문에 제안하는 ViT와 유사도 기반의 모델이 우수하다고 판단하기에는 어려움이 있다.

### 3-2. 시사점 및 한계점

먼저, 시사점은 기존 동일한 주제 및 컨셉을 가진 연구보다 Edge-Loc 관련 결함 유형에서 더 높은 수치의 예측률을 보여주었다. 해당 주제의 경우 CNN을 주로 사용하는데 Vision Transformer로 구현해 본 점에서 다른 방법론을 이용한 시도를 볼 수 있다.

하지만 한계점은 특정 결함은 어느 정도 탐지하지만 대다수의 결함 유형을 탐지하는 데 있어서 좋은 성능을 보이지 못했다. 또한, Model 구현에 있어서 파라미터들의 최적값을 찾지 못했다. 결함 유형 탐지에 있어서 특히, Scratch와 Edge\_Loc에서 가장 낮은 수치가 나왔는데, 해당 두 결함 유형은 타 결함에 비해 굉장히 두껍게 형성되지 않고 얇게 형성되는 데이터임을 확인했다. 이를 통해 얇게 형성되는 결함 유형 탐지에는 부적합하다고 생각한다. 또한, KNN을 통해 최종적으로 데이터의 결함 유형을 예측할 때, 서로 다른 Label를 2개만 예측하도록 설정했다. 이는 이중 유형의 결함을 예측하는데 있어서는 유용할 수 있지만 그 이상의 다중 유형의 결함에 대한 탐지에 있어서는 단점이 된다.

향후 연구 방향으로서는 제안한 프레임워크는 최종적으로 투영된 96차원의 임베딩 값을 가지고 Euclidean 또는 Cosine 유사도 계산을 진행하였는데 해당 부분에서 사용한 유사도의 경우에는 일반적으로

저차원에서 많이 사용하는 계산 방식이다. 이러한 부분이 모순적으로 적용되는 것을 통해 96차원의 Embedding 값을 PCA나 t-SNE를 통해 2차원 또는 3차원으로 축소한 뒤 유사도 계산을 통해 혼합 결함 데이터에 대한 유형을 예측하면 더 좋은 수치를 도출할 수 있을 가능성이 높다고 생각한다. 또한, 해당 연구의 주요 주제는 학습되지 않는 데이터에 대한 예측이다. 이러한 부분에서 학습 데이터를 이용하여 가상 데이터를 생성하고 이에 대해 모델에 학습하는 방식이 AI, 기계학습, 딥러닝에 좀 더 적합하다는 것을 느꼈다. 그리하여 GAN(Generative Adversarial Networks)를 활용하여 혼합 유형의 결함 데이터를 가상으로 생성한 뒤 모델에 학습시켜 분류하는 방법도 고려해 볼 만하다는 것을 생각해 보게 되었다. 또 다른 방식은 엔지니어가 새로운 결함 데이터를 보았을 때, 해당 엔지니어의 의견이나 진단 내용을 기반으로 NLP 기반의 Zero-Shot Learning도 추후 연구로 좋은 방법론이 될 수 있겠다는 생각이 들었다.

해당 연구 주제는 비록 좋은 결과를 도출하지 못했지만, 굉장히 중요한 주제라고 생각한다. 반도체 뿐만 아니라 제조 분야에서 새로운 결함이 등장하는 것은 기술의 발전에 따라 발생하는 Trade-Off 관계라고 생각한다. 따라서, 새로운 결함이나 혼합 결함을 정확하게 판단하고 예측하는 관련 주제를 많은 연구자들이 연구하고 방법을 창출한다면 앞으로의 제조업은 더욱더 비용·시간적인 부분에서 많은 부분을 현재보다 절약하고 손실을 최대한 막을 수 있을 것으로 생각한다.

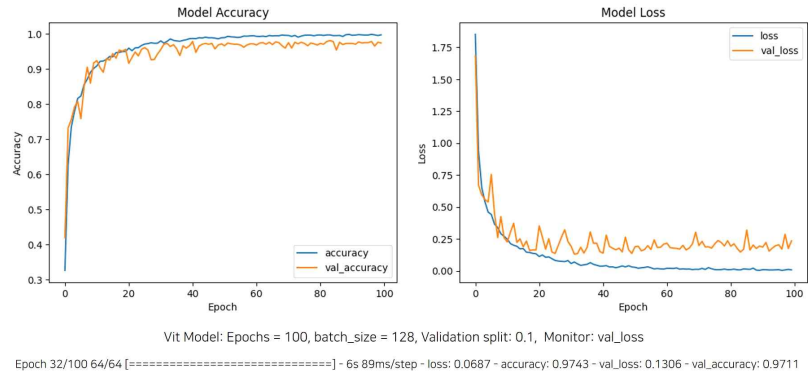


Fig. 9. 정상과 단일 유형의 결함 데이터를 9:1로 학습시킨 ViT 모델

평가지표	Vit + Cosine	Vit + Euclidean
Accuracy	21.4692%	25.4154%
F1-Score	47.2479%	48.6249%
Recall	46.8894%	47.6562%
Precision	53.5908%	55.6471%

Label	유형	Vit + Cosine	Vit + Euclidean
Center+Edge_Loc	[1 0 1 0 0 0 0]	10.6%	61.6%
Center+Edge_Ring	[1 0 0 1 0 0 0]	12.5%	0.0%
Center+Loc	[1 0 0 0 1 0 0]	38.4%	40.4%
Center+Scratch	[1 0 0 0 0 1 0]	12.9%	16.8%
Donut+Edge_Loc	[0 1 1 0 0 0 0]	4.5%	75.0%
Donut+Edge_Ring	[0 1 0 1 0 0 0]	19.6%	0.0%
Donut+Loc	[0 1 0 0 1 0 0]	6.1%	52.0%
Donut+Scratch	[0 1 0 0 0 1 0]	8.0%	1.0%
Edge_Loc+Loc	[0 0 1 0 1 0 0]	52.1%	0.0%
Edge_Loc+Scratch	[0 0 1 0 0 1 0]	34.5%	50.0%
Edge_Ring+Loc	[0 0 0 1 1 0 0]	1.0%	0.0%
Edge_Ring+Scratch	[0 0 0 1 0 1 0]	0.0%	0.0%
Loc+Scratch	[0 0 0 0 1 0 1]	87%	83.7%

Fig. 10. 혼합(이중) 유형의 결함인 테스트 데이터에 대한 성능지표 결과

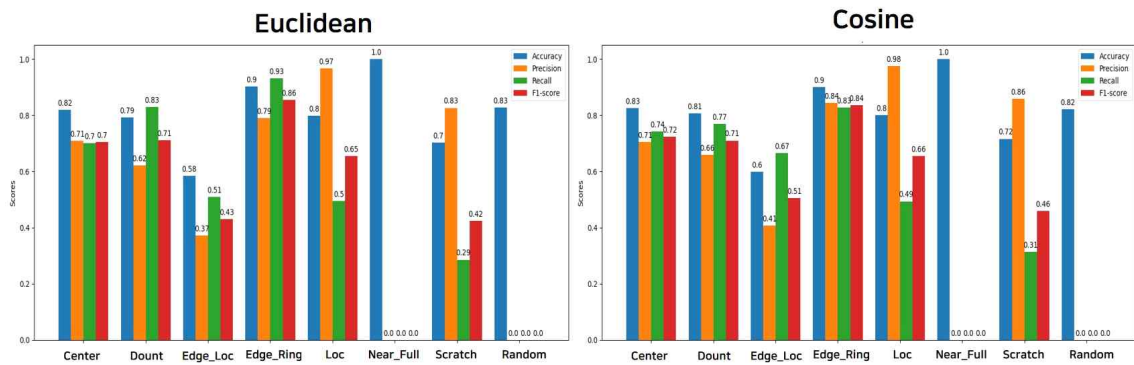


Fig. 11. ViT 모델과 각 유사도 기반 단일 유형의 결함 데이터에 대한 성능지표 결과

Label	유형	Vit + Cosine	Vit + Euclidean	선행연구 (CNN + Convolutional Autoencoder + 임계값)
Center+Edge_Loc	[1 0 1 0 0 0 0 0]	10.6%	61.6%	59%
Center+Edge_Ring	[1 0 0 1 0 0 0 0]	12.5%	0%	69%
Center+Scratch	[1 0 0 0 0 0 1 0]	12.9%	16.8%	40%
Edge_Loc+Scratch	[0 0 1 0 0 0 1 0]	34.5%	50.0%	41.6%
Edge_Ring+Scratch	[0 0 0 1 0 0 1 0]	0%	0%	64.4%

Fig. 12. Byun, Yunseon and Baek, Jun-Geol (2020)와 결과값 비교

## Reference

- [1] Chien, C. F., Hsu, S. C., & Chen, Y. J. (2013). A system for online detection and classification of wafer bin map defect patterns for manufacturing intelligence. *International Journal of Production Research*, 51(8), 2324-2338.
- [2] Jeong, Y. S., Kim, S. J., & Jeong, M. K. (2008). Automatic identification of defect patterns in semiconductor wafer maps using spatial correlogram and dynamic time warping. *IEEE Transactions on Semiconductor manufacturing*, 21(4), 625-637.
- [3] Kim, J., Lee, Y., & Kim, H. (2018). Detection and clustering of mixed-type defect patterns in wafer bin maps. *Iise Transactions*, 50(2), 99-111.
- [4] Kim, H. K., & Shim, J. (2024). Generalized zero-shot learning for classifying unseen wafer map patterns. *Engineering Applications of Artificial Intelligence*, 133, 108476.
- [5] Byun, Y., & Baek, J. G. (2020, January). Mixed Pattern Recognition Methodology on Wafer Maps with Pre-trained Convolutional Neural Networks. In *ICAART* (2) (pp. 974-979).
- [6] Shim, J., & Kang, S. (2023). Learning from single-defect wafer maps to classify mixed-defect wafer maps. *Expert Systems with Applications*, 233, 120923.
- [7] Yu, N., Chen, H., Xu, Q., & Hasan, M. M. (2021, October). Wafer map defect recognition with few shot learning based on hybrid self-attention mechanism and prototype network. In *2021 China Automation Congress (CAC)* (pp. 4128-4134). IEEE.
- [8] Wang, R., & Wang, S. (2023). Similarity searching for fault diagnosis of defect patterns in wafer bin maps. *Computers & Industrial Engineering*, 185, 109679.
- [9] Zhao, J., & Tang, L. (2023, June). Metavit-Trans: A Framework for Mixed-Type Defect Detection of Wafers with Vision Transformer Combined with Meta-Learning and Transfer Learning. In *2023 China Semiconductor Technology International Conference (CSTIC)* (pp. 1-3). IEEE.
- [10] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., ... & Polosukhin, I. (2017). Attention is all you need. *Advances in neural information processing systems*, 30.
- [11] Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., ... & Housby, N. (2020). An image is worth 16x16 words: Transformers for image recognition at scale. *arXiv preprint arXiv:2010.11929*.

## Authors

- Jae-June Choe: Undergrate Student., Dept. of Industrial Engineering, Kumoh National Institute of Technology, Korea; [cjj1222@kumoh.ac.kr](mailto:cjj1222@kumoh.ac.kr)
- Seong-Ju Cho: Ph.D., professor, Dept. of Industrial Engineering, Kumoh National Institute of Technology, 61 Dachak-ro (yangho-dong), Gumi, Gyeongbuk 39177, Korea; [csj@kumoh.ac.kr](mailto:csj@kumoh.ac.kr)