

Multi-class Data Description 기반의 웨이퍼 빈 맵 불량 패턴 분류 및 신규 불량 패턴 검출방법

송창용¹ · 정영선² · 김병훈^{1*}

¹한양대학교 산업경영공학과 / ²전남대학교 산업공학과 및 아트&디자인 테크놀로지 협동과정

A Multi-Class Data Description Based Method for Classifying Predefined Defect Patterns and Detecting New Defect Patterns of Wafer Bin Maps

Changyong Song¹ · Young-Seon Jeong² · Byunghoon Kim¹

¹Department of Industrial and Management Engineering, Hanyang University

²Department of Industrial Engineering and Interdisciplinary Program of Arts&Design Technology, Chonnam National University

In the semiconductor manufacturing process, electrical die sorting(EDS) test is a process to check whether the individual chips have reached the desired quality level by performing electrical property inspection on each semiconductor chip on a wafer. After performing the EDS test, a human expert can check the spatial defect pattern of wafer bin map(WBM) that displays the defectiveness of each chip. Since the defect patterns of WBMs are analyzed to trace back the processes and facilities for root cause analysis, it is very important to classify the WBM defect patterns. Because most existing studies use a classification model that trains predefined defect patterns, there is a limitation in that it cannot detect a new defect pattern that is not included in the training dataset. In this study, a method for classifying predefined defect patterns and detecting undefined defect patterns is proposed based on the multi-class data description model. As a result of the evaluation using actual WBM data, it was confirmed that it has excellent performance in classifying existing defect patterns and detecting new defect patterns.

Keywords: Semiconductor, Wafer Bin Map, Defect Pattern Classification, New Defect Pattern Detection, Multi-Class Data Description

1. 서론

반도체 제조 공정은 산화, 포토, 에치 등의 단위 공정을 기반으로 다양한 단위 공정 조합으로 이루어져 있으며, 웨이퍼 투입 이후 반도체가 완성되기 위해서는 약 3~4개월 정도의 시간이 걸린다 (Yoo *et al.*, 2015). 최근에는 공정의 미세화, 다양화와 같은 이유로

공정의 난이도가 증가했다. 이에 따라 수율 향상을 위해 불량 검출 및 원인 파악이 반도체 제조 공정의 핵심 요소 중 하나로 꼽힌다 (Jin *et al.*, 2019). 반도체 제조 공정 중 electrical die sorting(EDS) 테스트는 웨이퍼에 형성되어 있는 각 반도체 칩들에 대해 전기적 특성 검사를 실시해 각 칩들이 원하는 품질 수준에 도달했는지 확인하는 공정이다. <Figure 1>과 같이 EDS 테스트를 하고 나면

이 논문은 산업통상자원부 및 한국산업기술진흥원의 재원으로 한국전자기술연구원의 지원을 받아 수행된 인력양성사업임(P0008691). 한국연구재단의 연구비 지원을 받아 수행됨(No. NRF-2019R1F1A1042307).

* 연락처 : 김병훈 교수, 15588 경기도 안산시 상록구 한양대화로 55 5공학관 510호, 한양대학교 산업경영공학과, Tel : 031-400-5269,

Fax : 031-400-5959, E-mail : byungkim@hanyang.ac.kr

2021년 12월 30일 접수; 2022년 2월 16일, 2022년 4월 4일 수정본 접수; 2022년 3월 7일 게재 확정.

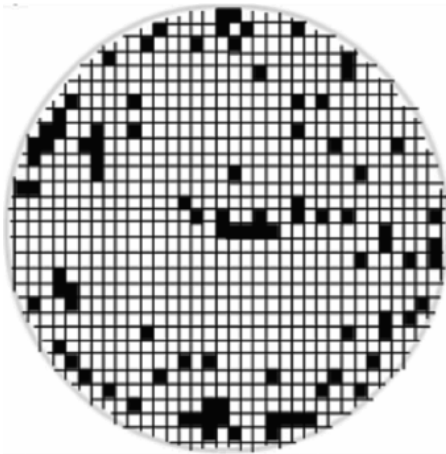


Figure 1. An Example of Wafer Bin Map

각 칩의 양품과 불량을 표시한 wafer bin map(WBM)을 확인할 수 있다(Liao *et al.*, 2013). WBM에는 각 칩의 불량 여부와 위치 정보가 같이 표시되기 때문에 특정 지역에서 많은 불량이 발생할 경우 같은 불량 패턴을 가지는 것으로 판단한다.

같은 불량 패턴을 가지는 WBM의 경우, 특정 공정 및 설비의 영향을 받아 불량률의 원인이 동일할 가능성이 크다(Jeong *et al.*, 2008). 예를 들어, 웨이퍼 가장자리 주위의 고리 모양 패턴은 예칭 공정의 문제일 가능성이 높고, 원형 및 스크래치 패턴은 chemical-mechanical polishing(CMP) 공정의 문제와 관련이 있다(Kang, 2020).

현업 엔지니어들은 이러한 근거를 바탕으로 하여 불량률의 원인을 찾아내기 때문에 WBM의 불량 패턴을 정확하고 빠르게 분류하는 것이 중요하다. 그러나 대부분의 기업에서 현업 엔지니어의 역량에 의존하여 육안으로 WBM의 불량 패턴을 분류하고 있다(Liu *et al.*, 2002). 이와 같이 엔지니어가 웨이퍼를 한 장씩 직접 판단하는 것은 시간과 비용이 많이 소요되어 비효율적이다(Kim *et al.*, 2020). 또한 엔지니어의 지식과 경험에 따라 동일한 불량 패턴을 가진 웨이퍼를 다른 불량 패턴으로 판단할 수 있다(Chen *et al.*, 2000). 따라서 높은 정확도로 웨이퍼의 불량을 분류하는 기계학습 모델을 개발하는 연구는 매우 중요하다.

대부분의 기존 연구에서는 분류 모델을 사용하여 WBM의 불량 패턴을 분류하기 때문에 사전에 정의된 불량 패턴의 분류에 대해서는 높은 정확도를 보이지만 신규 불량 패턴을 탐지하지 못한다는 한계가 존재한다. 제조 현장에서 신규 불량 패턴을 탐지하지 못한다면 불량률의 원인이 되는 공정 및 설비를 추적하는 단계에서 어려움을 겪을 수 있다. 이로 인해 불량률의 원인을 정확하게 파악하지 못해 재작업을 해야 할 수도 있다. 따라서 신규 불량 패턴을 탐지할 수 있다면 불량률의 원인을 찾는 단계에서 시간과 비용을 절약할 수 있다.

이러한 문제를 해결하기 위해 본 연구에서는 사전에 정의된 불량 패턴을 분류함과 동시에 신규 불량 패턴을 탐지하는 방법을 제안한다. 본 연구의 전체 프레임워크는 다음과 같이 구성되어 있다. 불량 패턴을 보다 명확하게 표현하기 위해 WBM

의 노이즈로 판단되는 칩을 제거하고, 합성곱 신경망 모델을 이용해 사전에 정의된 불량 패턴을 표현할 수 있는 특징을 추출한다. 그 후 Deep-MCDD(Lee *et al.*, 2020) 기반의 multi-class data description 모델을 이용해 합성곱 신경망을 통해 추출한 사전에 정의된 불량 패턴과 신규 불량 패턴의 특징을 구분하는 방법을 제안한다. 추출한 사전에 정의된 불량 패턴의 특징들을 각 불량 패턴별로 독립된 영역을 이루도록 학습하고, 각 샘플과 영역 사이의 거리를 이용하여 사전에 정의된 불량 패턴을 분류함과 동시에 신규 불량 패턴을 탐지한다.

본 논문은 다음과 같이 구성되어 있다. 제2장에서는 웨이퍼 불량 패턴의 분류와 신규 불량 패턴 탐지와 관련한 선행 연구를 소개한다. 제3장에서는 노이즈 제거 및 특징 추출, multi-class data description 모델 학습과 임계값 설정 방법에 대해 서술한다. 제4장에서는 데이터 소개 및 제안된 방법의 성능을 평가한다. 제5장에서는 본 논문의 결론과 기대 효과에 대해 논의한다.

2. 선행 연구

2.1 WBM 불량 패턴 분류

기존 연구에서는 다양한 방법을 WBM 불량 패턴 분류에 이용하였다. Li and Huang(2009)은 self-organizing map(SOM)과 support vector machines(SVM)를 결합한 방식을 WBM 불량 패턴 분류에 이용한다. 각 WBM의 특징을 추출한 후 SOM을 이용하여 군집화하고, SVM을 이용하여 WBM의 불량 패턴을 분류한다. 하지만 SVM 모델의 특성상 적절한 하이퍼 파라미터를 찾기 위해 많은 시간이 소요된다는 단점이 있다. Piao *et al.*(2018)은 중앙값 필터를 이용해 노이즈의 영향을 줄이고, 라돈 변환 후 평균, 최댓값, 최솟값, 표준 편차와 같은 특징을 추출하고, 의사결정나무를 기반으로 한 앙상블 기법을 이용해 불량 패턴을 분류한다. 웨이퍼 사이즈에 관계없이 적용할 수 있다는 장점이 있지만, 추출할 수 있는 특징의 수가 한계가 있기 때문에 scratch 등 일부 패턴에 대해 분류 성능이 떨어지는 한계가 있다. Saqlain *et al.*(2019)은 각 웨이퍼에서 density, geometry, radon 기반의 특징을 추출한 후, 추출한 특징들을 logistic regression, random forests, gradient boosting machine, artificial neural network 분류기에 학습한다. 그 후, 각 분류기의 정확도를 기반으로 가중치를 부여하는 soft voting ensemble(SVE) 기법을 이용하여 불량 패턴을 분류하였다. 하지만 다양한 특징 추출과 분류기 학습을 위해 많은 시간이 들어간다는 한계가 있다.

최근에는 이미지 패턴 인식 분야에서 합성곱 연산을 통해 이미지의 특징을 추출하는 합성곱 신경망 모델을 많이 사용하고 있다. 따라서 합성곱 신경망 모델을 이용해 웨이퍼 불량 패턴을 분류하는 연구도 활발히 진행되고 있다. Nakazawa *et al.*(2018)은 사전에 정의된 22개의 불량 패턴에 대해 시뮬레이션으로 웨이퍼 맵

을 만들었고, 이를 합성곱 신경망 모델로 학습하여 불량 패턴 분류에 사용했다. 하지만 실제 공정 데이터를 이용해 평가한 결과 curved scratch 등 일부 불량 패턴에 대해 낮은 정확도를 보였다. Wang *et al.*(2019)은 각 불량 패턴의 데이터 개수에 따른 데이터 불균형으로 인해 학습 능력이 떨어지는 현상을 generative adversarial networks(GAN)을 이용해 불충분한 클래스의 데이터를 생성함으로써 전체 분류 정확도를 향상하였다. 하지만 많은 훈련 샘플을 생성 및 학습하기 때문에 SVM, convolutional neural networks(CNN) 등 다른 분류 모델에 비해 학습 시간이 길다는 한계가 있다. Saqlain *et al.*(2020)은 WBM의 각 불량 패턴의 클래스 불균형을 rotation, shift 등의 data augmentation 기법을 활용하여 해결함으로써 분류 성능을 향상하였다. 그러나 이러한 연구들은 신규 불량 패턴의 발생을 가정하지 않는 분류 모델을 사용하기 때문에 사전에 정의된 불량 패턴에 대한 분류에서는 높은 정확도를 보이지만, 신규 불량 패턴을 탐지하지 못한다는 한계를 가지고 있다.

2.2 WBM 신규 불량 패턴 탐지

Baek *et al.*(2020)은 기존 불량 패턴을 학습한 합성곱 신경망 모델에 의해 계산된 소프트맥스 확률값을 이용해 신규 불량 패턴을 탐지하였다. 합성곱 신경망 모델에 의해 계산된 소프트맥스 확률값의 최댓값과 표준편차를 기반으로 임계값을 설정하고, 임계값을 넘지 못하면 신규 불량 패턴을 탐지하였다. 하지만 합성곱 신경망 기반의 분류 모델은 신규 불량 패턴에 대해 과신하는 경향이 있어 정확한 확률 정보를 제공하지 않는다. 따라서 신규 불량 패턴의 탐지 성능이 떨어진다는 한계가 있다. Jang *et al.*(2020)은 합성곱 신경망 기반의 one-versus-one(OVO) 앙상블 모델을 이용해 신규 불량 패턴을 탐지했다. 특징 공간에서 각 불량 패턴의 근접도를 기반으로 가중 평균 점수를 계산해 불량 패턴의 유무를 판단하고, 모든 불량 패턴이 존재하지 않으면 신규 불량 패턴으로 판단한다. 하지만 OVO 기반의 분류기를 사용하므로 기존 불량 패턴의 종류가 늘어날수록 학습해야 하는 분류기의 수가 늘어나 학습시간이 길어진다는 한계가 존재한다. Kim *et al.*(2021)은 기존 불량 패턴 분류와 신규 불량 패턴 탐지를 위해 학습 과정에서 기존 불량 패턴 데이터와 학습 데이터로부터 멀리 떨어진 데이터, 즉

out-of-distribution(OOD) 데이터를 모두 활용한다. 기존 불량 패턴 데이터는 정답 클래스에 속할 확률이 높게, OOD 데이터는 모든 클래스에 대하여 확률이 동일하게 나올 수 있도록 학습한다. 그 후 임계값을 설정해 각 샘플의 확률값이 임계값을 넘지 못하면 신규 불량 패턴으로 판단한다. 하지만 실제 제조 현장에서는 사전에 정의된 불량 패턴 이외에 신규 불량 패턴이 발생하는 경우가 많지 않아 학습에 필요한 충분한 수의 OOD 데이터를 수집하기 어렵다는 한계가 있다.

3. 제안 방법론

본 연구의 목적은 WBM의 사전에 정의한 불량 패턴을 분류함과 동시에 신규 불량 패턴을 탐지하는 것이다. 따라서 먼저 불량 패턴을 보다 명확하게 표현하기 위해 WBM의 노이즈로 판단되는 칩을 제거한 후, 합성곱 신경망 모델을 이용해 각 불량 패턴이 구분되어 표현될 수 있는 특징을 추출한다. 그 후 Deep-MCDD(Lee *et al.*, 2020) 기반의 multi-class data description 모델을 통해 사전에 정의된 불량 패턴의 특징들을 불량 패턴별로 독립된 영역을 이루도록 학습하고, 각 샘플과 가장 가까운 영역과의 거리를 이용해 사전에 정의된 불량 패턴을 분류함과 동시에 신규 불량 패턴을 탐지한다.

제안하는 방법의 구조는 <Figure 2>와 같이 4단계로 구성된다. 첫 번째 단계는 WBM의 불량 패턴을 보다 명확하게 표현하기 위해 노이즈로 판단되는 불량 칩을 제거한다. 불량 칩의 위치에 따라 다른 필터를 적용하였고 제3.1절에서 해당 필터에 대해 설명한다. 두 번째 단계는 각 불량 패턴의 특징 추출 단계이다. 합성곱 신경망의 특징 추출기 부분을 이용하여 각 불량 패턴이 구분되어 표현될 수 있는 특징을 추출하였고 제3.2절에서 해당 구조에 대해 설명한다. 세 번째 단계는 Deep-MCDD(Lee *et al.*, 2020) 기반의 multi-class data description 모델을 사용하여 합성곱 신경망을 이용해 추출한 사전에 정의된 불량 패턴의 특징들이 각 불량 패턴별로 독립된 영역을 이룰 수 있도록 학습하였고, 제3.3절에서 해당 모델에 대해 설명한다. 마지막 단계는 사전에 정의한 불량 패턴을 분류함과 동시에 신규 불량 패턴을 탐지하는 단계이다. 각 샘플과 가장 가까운 영역까지의 거리를 $S(\mathbf{x})$ 로 정의하고, $S(\mathbf{x})$ 의 분포를 추정하여 임계값을 설정한

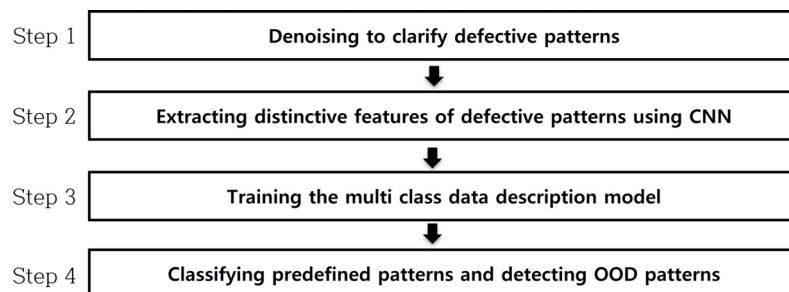


Figure 2. Full Framework of the Proposed Method

후, 임계값을 넘으면 신규 불량 패턴으로 판단하며 제3.4절에서 설명한다.

3.1 불량 패턴 표현을 위한 노이즈 제거

제3.1절에서는 WBM에 존재하는 불량 패턴을 명확하게 표현하기 위해 필터를 사용해 노이즈로 판단되는 칩을 제거하는 방법에 대해 설명한다. 웨이퍼의 결함 패턴은 일반적으로 global defects와 local defects 두 가지로 분류된다. global defects는 클린룸의 입자와 같은 무작위 원인으로 인해 웨이퍼 전체에 흩어져 있는 결함을 의미한다. 반면 local defects는 결함들이 군집을 이루고 있는 형태를 보이며, 이를 통해 불량 원인이 되는 공정 및 설비를 추적할 수 있다(Yuan *et al.*, 2011). 본 연구에서는 WBM에서 군집을 이루지 않아 불량 원인을 추적하는데 영향을 주지 않는 global defects를 노이즈로 판단하고, 노이즈로 판단되는 칩은 제거하여 분석에서 제외한다.

$l \times l$ 크기의 필터를 만든 후 필터의 격자에 동일한 가중치를 부여하고, WBM을 구성하고 있는 수와 가중치를 곱한 합이 임계값보다 작은 경우에 노이즈로 판단하여 분석에서 제외한다. 하지만 불량 칩의 위치에 관계없이 같은 필터를 적용할 경우 경계면에 있는 칩은 경계면에 있지 않은 칩보다 주위에 있는 칩의 수가 적어 필터를 적용할 수 있는 칩의 개수가 적다. 이로 인해 WBM의 구성 값과 가중치를 곱한 합이 작아져 경계면에서 군집을 이루고 있는 불량 칩이 제거되는 문제가 발생한다. 따라서 불량 칩이 경계면에 있는 경우와 아닌 경우에 주위에 있는 칩의 개수를 고려하여 다른 가중치 계산 방법을 적용하였다(Kim *et al.*, 2015).

식 (1)은 불량 칩이 WBM의 경계면에 있지 않은 경우의 가중치 합을 구하는 식으로 불량 칩에 대하여 필터를 적용하고, 가중치와 WBM의 구성 값을 곱한 합이 임계값보다 작을 경우에 노이즈로 판단한다. 식 (1)에서 l 은 필터의 크기, $t = \frac{l-1}{2}$, $x_{i,j}$ 는 i, j 좌표에서 WBM을 구성하고 있는 값이다. $x_{i,j}$ 는 i, j 좌표에 정상 칩이 존재할 경우 1의 값을, 불량 칩이 존재할 경

우에는 2의 값을 가진다.

$$R(i,j) = \frac{1}{l^2} \sum_{m=-t}^t \sum_{n=-t}^t (x_{i,j} - 1)(x_{i+m,j+n} - 1) \quad (1)$$

식 (2)는 불량 칩이 경계면에 위치하고 있을 경우 가중치 합을 구하는 식이다. 구한 값이 임계값보다 작은 경우에 노이즈로 판단한다. 식 (2)에서 l 은 필터의 크기, $t = \frac{l-1}{2}$, $N(i,j)$ 은 해당 불량 칩을 둘러싸고 있는 칩의 개수이고, $x_{i,j}$ 는 i, j 좌표에서 WBM을 구성하고 있는 값이다. $A(i,j)$ 는 WBM의 i, j 좌표에 칩이 존재할 경우는 1의 값을, 존재하지 않을 경우에는 0의 값을 가진다. <Figure 3>은 경계면에 위치한 경우와 아닌 경우에 3×3 필터를 적용한 예시이다. 경계면에 위치하지 않은 경우는 각 격자에 $1/9$ 의 가중치가 부여되고, 경계면에 위치한 경우에는 각 격자에 $1/6$ 의 가중치가 부여된다.

$$R(i,j) = \frac{1}{N(i,j)} \sum_{m=-t}^t \sum_{n=-t}^t (x_{i,j} - 1)(x_{i+m,j+n} - 1)A(i+m,j+n) \quad (2)$$

3.2 합성곱 신경망을 이용한 각 불량 패턴의 특징 추출

제3.2절에서는 합성곱 신경망의 특징 추출기 부분을 이용해 각 불량 패턴의 구분되는 특징 추출 과정을 설명한다. 본 연구에서는 사전에 정의한 불량 패턴을 합성곱 신경망 모델에 학습한 후, 특징 추출기 부분을 이용해 각 불량 패턴의 특징을 추출한다. 각 WBM이 합성곱 신경망의 특징 추출기 부분을 통과하게 되면 각 WBM의 불량 패턴별로 구분되는 특징을 얻을 수 있다. 사전에 정의한 불량 패턴 샘플이 특징 추출기를 통과하면 사전에 정의한 불량 패턴끼리 서로 구분할 수 있는 특징을 얻을 수 있고, 신규 불량 패턴 샘플이 특징 추출기를 통과하게 되면 기존 불량 패턴과 구분되는 특징을 얻을 수 있다.

합성곱 신경망은 합성곱 연산을 통해 특징을 추출해 학습하

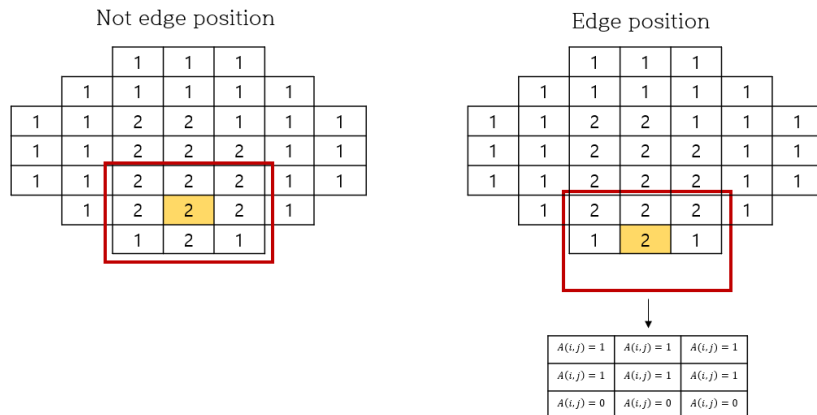


Figure 3. Examples of Applying 3x3 Filter According to Chip Position

는 네트워크 구조로 주로, 음성, 이미지 처리 분야에서 많이 사용되고 있다. 합성곱 신경망은 입력, 출력, 은닉층의 구조로 구성되어 있다. 데이터가 입력으로 들어오면 은닉층 안에서 합성곱, 풀링과 같은 연산을 반복하여 특징 맵을 얻어낸다. 이러한 과정을 통해 특징을 추출하고, 추출한 특징을 이용해 출력층에서 클래스를 분류한다. <Figure 4>는 본 연구에서 사용한 합성곱 신경망 모델 구조로 6개의 합성곱 층으로 구성되며 각각 64, 64, 128, 128, 256, 256채널을 가진다. 각 층마다 활성화 함수로는 rectified linear unit(RELU) 함수를 사용하였고, 2*2 맥스 풀링을 사용해 크기를 반으로 줄였다. 합성곱, 풀링 등의 연산을 모두 마친 후 1024, 512, 256, 128개의 노드를 가지는 완전 연결층을 통해 다시 연결하였고, 마지막 출력층은 사전에 정의한 불량 패턴의 클래스의 수와 같은 노드 수로 구성한다. 이 중 출력층을 제외한 부분을 특징 추출기로 사용하며 각 WBM의 특징으로 128개의 차원으로 추출할 수 있도록 구성하였다.

3.3 Multi-class Data Description 모델 학습

제3.3절에서는 제3.2절에서 추출한 사전에 정의한 불량 패턴의 특징을 입력으로 하여 Deep-MCDD 기반의 multi-class data description 모델을 학습한다. 이를 통해 사전에 정의한 불량 패턴의 특징들을 각 불량 패턴별로 독립된 영역을 이루도록 함으로써, 평가 데이터가 들어왔을 때 사전에 정의된 불량 패턴의 분류와 신규 불량 패턴을 탐지하는 과정을 설명한다. Deep-MCDD 기반의 multi-class data description 모델은 <Figure

5>와 같이 사전에 정의된 불량 패턴의 특징들을 각 클래스별로 특징 공간에서 가우시안 분포를 따르면서, 분리되어 있는 독립된 영역을 이루도록 학습한다. 형성된 영역을 바탕으로 각 샘플과 가장 가까운 클래스의 영역까지의 거리를 이용해 사전에 정의한 불량 패턴을 분류함과 동시에 신규 불량 패턴이 발생했을 시 탐지하고자 한다(Lee et al., 2020).

Deep-MCDD 모델은 각 클래스를 독립된 영역으로 학습하기 위해 두 가지 가정을 기반으로 한다. 첫째, 각 클래스에 속하는 샘플들은 해당 클래스의 조건부 확률 분포에 속해 있고, 다른 클래스의 분포와 완전히 분리되어 있다. 둘째, 각 클래스의 조건부 확률 분포는 특징 공간에서 각 클래스 k 의 평균 μ_k , 표준편차 σ_k 를 가지는 가우시안 분포를 따른다. 두 가정을 기반으로 하여 각 클래스 k 의 평균 μ_k 를 중심으로 하고 표준편차 σ_k 를 따르는 영역을 형성한다. 학습한 영역을 바탕으로 각 영역에 샘플이 속할 가능성을 나타내는 해당 클래스의 조건부 확률을 구하고, 이를 이용하여 식 (3)과 같이 거리 함수 $D_k(\cdot)$ 를 정의한다. 식 (3)에서 \mathbf{x} 는 입력 샘플, y 는 클래스, $f(\mathbf{x}; \mathbf{W})$ 는 Deep-MCDD 모델의 네트워크 매개변수 \mathbf{W} 를 가지는 인공신경망, d 는 특징 차원의 수를 의미한다.(Lee et al., 2020).

$$D_k(x) = -\log P(\mathbf{x}|y = k) = -\log N(f(\mathbf{x}; \mathbf{W}) | \mu_k, \sigma_k^2) \approx \frac{\|f(\mathbf{x}; \mathbf{W}) - \mu_k\|^2}{2\sigma_k^2} + \log \sigma_k^d \quad (3)$$

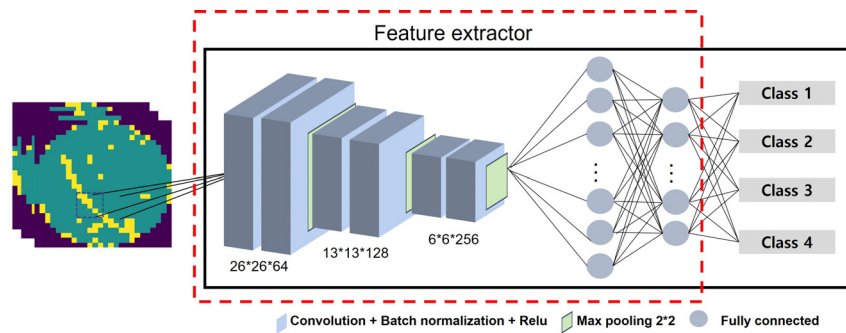


Figure 4. Structure of Convolution Neural Network Model

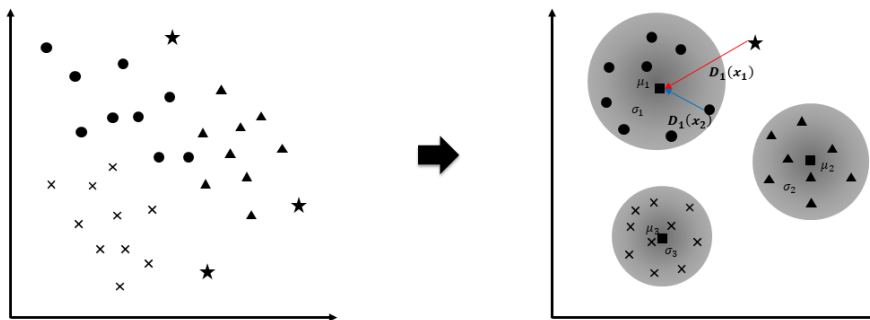


Figure 5. Training Multi-Class Data Description Model

학습 데이터의 분포를 특징 공간에서 각 클래스의 조건부 확률 분포와 일치하도록 학습하여, 각 클래스를 분류할 수 있도록 하는 최종 목적 함수는 식 (4)와 같다. 식 (4)의 첫 번째 항은 학습 데이터의 분포를 가우시안 분포와 일치하기 위한 항이고, 두 번째 항은 사전에 정의된 클래스를 분류하기 위한 항으로 정답 클래스에 대한 샘플의 사후 확률을 최대화한다. Deep-MCDD 모델은 각 불량 패턴의 특징들이 특징 공간에서 가우시안 분포를 따르도록 학습하기 때문에 gaussian discriminant analysis(GDA)를 적용할 수 있다. 따라서 각 클래스의 조건부 확률 분포 $P(\mathbf{x}|y)$ 과 사전 확률 $P(y)$ 를 이용하여 사후 확률 $P(y|\mathbf{x})$ 를 구할 수 있다. 각 클래스에 대한 사후 확률 $P(y=k|\mathbf{x})$ 은 $-D_k(\mathbf{x}) + b_k$ 로 표현되며, 각 클래스의 조건부 확률 분포 $P(\mathbf{x}|y=k)$ 과 사전 확률 $P(y=k)$ 을 이용해 구할 수 있다. 사후 확률 $P(y=k|\mathbf{x})$ 에서 b_k 는 각 클래스에 대한 바이어스 항을 의미한다. 식 (4)의 파라미터 v 는 두 항의 영향을 제어하는 정규화 계수를, K 는 클래스의 개수를 의미한다. 목적함수에 속해 있는 파라미터 $\mathbf{W}, \mu, \sigma, b$ 는 minibatch stochastic gradient descent(SGD)와 gradient back-propagation 방법을 사용하여 최적화한다(Lee et al., 2020).

$$\min_{\mathbf{W}, \mu, \sigma, b} \frac{1}{N} \sum_{i=1}^N \left[D_{y_i}(\mathbf{x}_i) - \frac{1}{v} \log \frac{\exp(-D_{y_i}(\mathbf{x}_i) + b_{y_i})}{\sum_{k=1}^K \exp(-D_k(\mathbf{x}_i) + b_k)} \right] \quad (4)$$

3.4 기존 불량 패턴 분류 및 신규 불량 패턴 탐지

식 (4)의 목적 함수는 학습 데이터가 속해 있는 클래스에 대한 사후 확률을 최대화하므로, 식 (5)와 같이 가장 높은 사후 확률을 가지는 클래스로 입력 샘플의 불량 패턴을 예측한다(Lee et al., 2020).

$$\begin{aligned} \hat{y} &= \operatorname{argmax}_k P(y=k|\mathbf{x}) \\ &= \operatorname{argmax}_k [-D_k(\mathbf{x}) + b_k] \end{aligned} \quad (5)$$

또한, 각 샘플과 특징 공간에서 가장 가까운 영역과의 거리를 이용하여 식 (6)과 같이 $S(\mathbf{x})$ 를 정의하고, 이를 이용하여 신규 불량 패턴을 탐지한다.

$$S(\mathbf{x}) = \min_k D_k(\mathbf{x}) \quad (6)$$

식 (6)에서 얻은 $S(\mathbf{x})$ 값은 각 샘플에서 가장 가까운 영역까지의 거리를 나타낸 것으로 $S(\mathbf{x})$ 값이 작을수록 해당 불량 패턴의 영역까지의 거리가 가깝고, $S(\mathbf{x})$ 값이 클수록 해당 불량 패턴의 영역까지의 거리가 멀다. 따라서 학습 데이터의 $S(\mathbf{x})$ 확률 분포를 추정하고 이를 이용해 신규 불량 패턴을 탐지한다. 본 연구에서는 kernel density estimation(KDE)방법을 이용해 학습 데이터의 $S(\mathbf{x})$ 확률 분포를 추정하고, 유의 수준을 0.05로 채택하여 상위 5%값을 임계값으로 설정하였다. 식 (7)은 KDE방법을 이용해 추정한 $S(\mathbf{x})$ 의 확률 밀도 함수이고, K 는 커널함수, h 는 커널함수의 bandwidth값으로 커널의 분산을 조절하는 파라미터, n 은 학습 데이터의 개수를 의미한다. 평균 데이터의 $S(\mathbf{x})$ 값이 임계값보다 작거나 같은 경우에는 사전에 정의한 불량 패턴으로 분류하고, 임계값보다 클 경우에는 신규 불량 패턴으로 분류한다. 본 연구에서는 가우시안 커널 함수와 h 값으로는 식 (8)을 통해 구한 값을 사용하였다(Higgins, 2004). 식(8)의 n 은 학습 데이터의 개수, $\hat{\sigma}$ 는 학습 데이터에서 추정한 $S(\mathbf{x}_i)$ 의 표준편차를 의미한다.

$$\hat{f}_h(S(\mathbf{x})) = \frac{1}{n^*h} \sum_{i=1}^n K\left(\frac{S(\mathbf{x}) - S(\mathbf{x}_i)}{h}\right) \quad (7)$$

$$h = \left(\frac{4\hat{\sigma}^5}{3n}\right)^{\frac{1}{5}} \approx 1.06\hat{\sigma}n^{-\frac{1}{5}} \quad (8)$$

4. 실험 결과

4.1 데이터 및 실험 설정

본 연구에서 사용한 WBM 데이터는 케글(<https://www.kaggle.com/qingyi/wm811k-wafer-map>)에 있는 오픈 데이터를 사용하였으며, 26*26 크기의 단일 종류의 WBM만을 사용하였다. 실험에 사용한 WBM은 <Figure 6>와 같이 표현되어 있으며 배경, 정상 칩, 불량 칩으로 구성되어 있다. 배경은 보라색을 띄고 있으며 0의 값을, 정상 칩은 초록색을 띄며 1의 값을, 마지막으로 불량 칩은 노란색을 띄며 2의 값을 가지고 있다. 본 실험에서는 <Figure 6>에서 보이는 center, edge-loc, loc, edge-ring, scratch, near-full, donut의 총 7개의 패턴을 사용하였다.

그러나 <Figure 7>에서 볼 수 있듯이 edge-ring 패턴 데이터와 edge-loc 패턴의 일부 데이터가 유사한 점이 많은 것을 확인

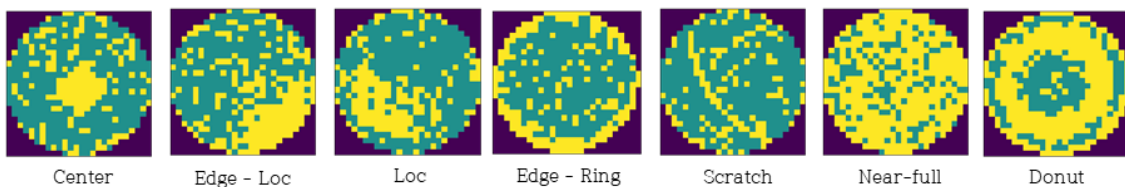


Figure 6. Examples of WBM Defect Patterns

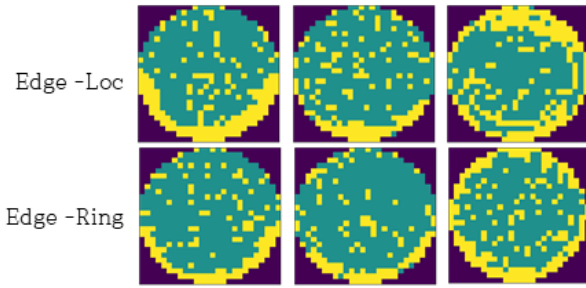


Figure 7. Examples of Edge-Loc Patterns And Edge-Ring Patterns

Table 1. The Number of Wafers for Each Defect Pattern Before and After Reclassification

Defect patterns	Before reclassification	After reclassification
	Number of wafers	Number of wafers
Center	90	90
Edge-Loc	296	296
Edge-Ring	31	-
Loc	297	328
Scratch	72	72
Near-full	16	16
Donut	1	1
Total	803	803

할 수 있다. 따라서 edge-ring 패턴 데이터를 edge-loc 패턴으로 재분류하였다.

<Table 1>에서 edge-ring 패턴을 재분류하기 전과 후의 WB M 불량 패턴의 데이터 개수를 확인 할 수 있다. 본 연구에서는 사전에 정의한 불량 패턴 분류와 신규 불량 패턴 탐지에 관하여 실험을 진행하므로, 신규 불량 패턴에 대한 가정이 필요하다. 신규 불량 패턴은 기존 불량 패턴에 비해 발생 빈도가 낮을 것이므로, 데이터 개수를 기반으로 center, edge-loc, loc, scratch 패턴은 사전에 정의한 불량 패턴으로, near-full, donut 패턴은 신규 불량 패턴으로 가정하였다.

본 연구의 학습 데이터 셋으로는 전체 사전에 정의된 불량 패턴 데이터 셋 중 80%를 이용하였고, 평가 데이터 셋은 20%의 사전에 정의된 불량 패턴 데이터와 신규 불량 패턴 데이터로 구성하였다. 적은 수의 데이터 셋으로 합성곱 신경망을 학습하기 위해 rotation, zoom, horizontal flipping, width shift, height shift, shear range 와 같은 데이터 증강 기법을 이용하여 추가적으로 학습 데이터를 생성해 학습에 이용하였다. 또한 데이

Table 2. Confusion Matrix of New Patterns Detection

		Actual Class	
		Predefined patterns	New patterns
Predicted Class	Predefined patterns	True Positive(TP)	False Negative(FN)
	New patterns	False Positive(FP)	True Negative(TN)

터 셋을 k 개의 동일한 크기로 분할하여 한 개의 겹은 검증 데이터로 나머지 $k-1$ 개의 겹은 학습데이터로 사용하고, 이를 반복하는 k -겹 교차 검증 방법을 활용하여 학습하였다. k 값은 5로 설정하였다. multi-class data description 모델의 데이터 셋은 합성곱 신경망의 특징 추출기를 통해 추출한 128차원의 각 불량 패턴의 특징을 이용하였다. 학습 데이터 셋은 학습을 위해 분할한 사전에 정의된 불량 패턴 데이터를 합성곱 신경망을 통해 추출한 특징으로 구성하였으며, 평가 데이터 셋은 평가를 위해 분할한 사전에 정의된 불량 패턴과 신규 불량 패턴의 특징으로 구성하였다.

본 연구에서 사용한 합성곱 신경망의 모델의 세부 설정은 다음과 같다. 학습률은 0.001, 배치의 크기는 32, epoch은 100으로 설정하였다. 또한 multi-class data description 모델의 구조와 세부 설정은 다음과 같다. 모델의 구조는 128개의 노드를 가지는 4층의 다층 퍼셉트론을 이용하였고, 각 층마다 활성화 함수로는 RELU를 사용하였다. 파라미터 v 는 0.01, 학습률은 0.001, 배치의 크기는 32, 최대 epoch은 100으로 설정하였다.

본 연구에서는 3가지 실험을 진행하였다. 첫 번째 실험은 원본 데이터와 노이즈로 판단되는 칩을 제거한 데이터의 성능을 비교하였다. 두 번째 실험은 본 연구에서 제안한 방법과 합성곱 신경망의 소프트맥스 확률값을 이용해 신규 불량 패턴을 탐지한 방법(Baek *et al.*, 2020)의 성능을 비교하였다. 세 번째 실험은 edge-ring 패턴의 재분류 전과 후의 성능 비교와 사전에 정의한 불량 패턴의 수를 달리하여 제안한 방법의 성능을 평가하였다.

4.2 성능 평가 지표

본 연구에서는 크게 사전에 정의한 불량 패턴 분류와 신규 불량 패턴 탐지에 대한 실험을 진행한다. 사전에 정의한 불량 패턴 분류를 위한 평가 지표로는 분류 정확도를 이용하며, 사전에 정의한 불량 패턴으로 가정한 데이터 중 모델이 정확하게 예측한 비율을 의미한다. 식 (9)에서 n_{ij} 는 불량 패턴 i 를 불량 패턴 j 로 예측한 데이터의 수를 의미한다.

$$Accuracy = \frac{\sum_{i=1}^k n_{ii}}{\sum_{i=1}^k \sum_{j=1}^k n_{ij}} \quad (9)$$

신규 불량 패턴 탐지에 대한 평가지표로는 사전에 정의한

불량 패턴을 true로, 신규 불량 패턴을 false로 간주하여 <Table 2>와 같이 true positive(TP), false negative(FN), true negative(TN), false positive(FP)를 사용하였다. TP는 실제 사전에 정의한 불량 패턴을 사전에 정의한 불량 패턴으로 예측한 경우의 비율이고, FP는 실제 사전에 정의한 불량 패턴을 신규 불량 패턴으로 예측한 비율이다. TN은 신규 불량 패턴을 신규 불량 패턴으로, FN은 신규 불량 패턴을 사전에 정의한 불량 패턴으로 예측한 비율이다.

4.3 Original WBM과 Denoised WBM 성능 비교

제4.3절에서는 원본 데이터와 노이즈로 판단되는 칩을 제거한 데이터의 성능을 비교하였다. 본 연구에서는 3*3필터와 0.4

의 임계값을 설정해 WBM의 노이즈로 판단되는 칩을 제거했다. <Figure 8>의 윗 행은 원본 데이터이고, 아래 행은 노이즈로 판단되는 칩을 제거한 데이터이다. 노이즈로 판단되는 칩을 제거한 데이터에서 더 뚜렷하게 각 불량 패턴의 특징을 확인할 수 있다.

<Figure 9(a)>은 사전에 정의한 불량 패턴의 원본 데이터와 노이즈로 판단되는 칩을 제거한 WBM의 예시이고, <Figure 9(b)>는 신규 불량 패턴으로 가정한 near-full 불량 패턴의 원본 데이터와 노이즈로 판단되는 칩을 제거한 WBM의 예시이다. <Figure 9>을 통해 노이즈로 판단되는 칩을 제거한 데이터가 신규 불량 패턴 데이터와 더 명확히 구분됨을 확인할 수 있다. 따라서 <Table 3>과 같이 원본 WBM과 노이즈로 판단되는 칩을 제거한 WBM 모두 사전에 정의한 불량 패턴의 분류에 대해

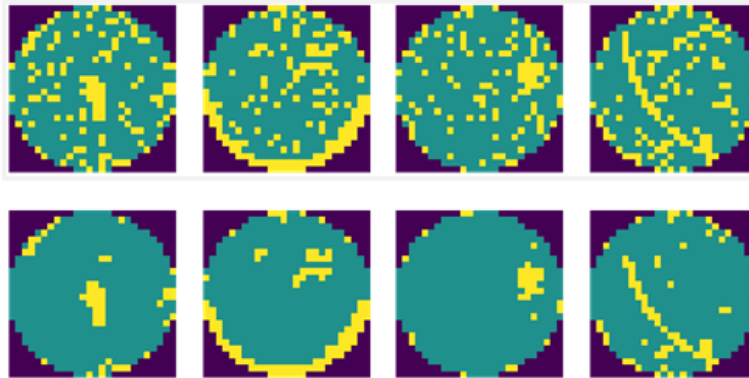


Figure 8. Examples of Original Data and Denoised Data

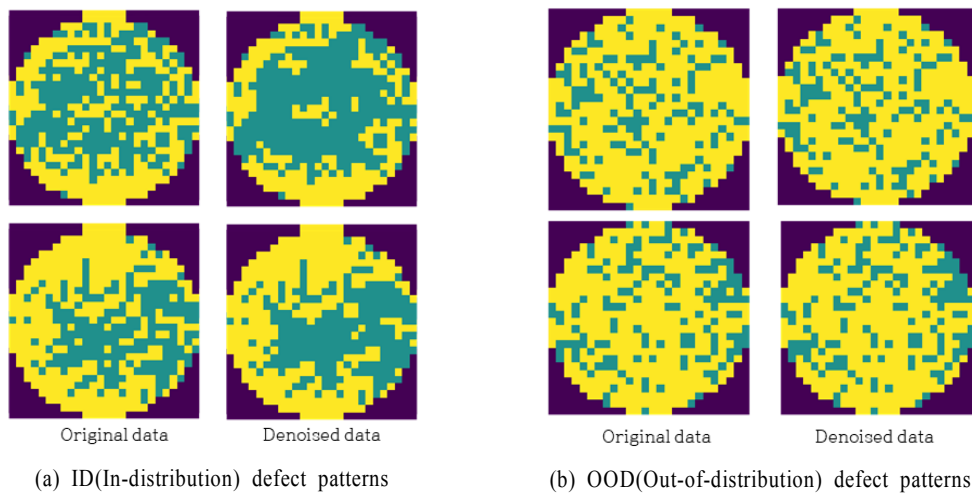


Figure 9. Examples of Original and Denoised Data of ID and OOD Defect Patterns

Table 3. Comparison of Experimental Results of Original Data and Denoised Data

	Accuracy(%)	TP(%)	FN(%)	TN(%)	FP(%)
Original data	92.4	94.9	5.1	64.7	35.3
Denoising data	94.3	96.2	3.8	94.1	5.9

서는 비슷한 성능을 보이지만, 노이즈로 판단되는 칩을 제거한 WBM이 기존 불량 패턴과 신규 불량 패턴을 더 명확하게 구분함으로써 신규 불량 패턴을 탐지함에 있어 더 높은 성능을 보임을 확인할 수 있다.

4.4 소프트맥스 확률 기반 모델과 성능 비교

제4.4절에서는 본 연구에서 제안한 방법과 합성곱 신경망의 소프트맥스 확률을 이용한 신규 불량 패턴 탐지에 관한 연구 (Baek et al., 2020)의 성능을 비교하였다. 식 (10)과 같이 검증 데이터에서 예측에 성공한 샘플 n 개가 있을 때, i 번째 샘플의 소프트맥스 확률의 최댓값을 M_i 라 하고, 샘플 n 개의 M_i 의 평균을 θ_{max} 라 정의한다. 또한 식 (11)과 같이 검증 데이터에서 예측에 실패한 표본 m 개가 있을 때, j 번째 샘플의 소프트맥스 확률의 표준 편차를 S_j 라 하고, 샘플 m 개의 S_j 의 평균을 θ_{std} 라 정의한다. 각 θ_{max} , θ_{std} 임계값으로 설정하고, 임계값을 넘지 못했을 경우 신규 불량 패턴으로 판단한다.

$$\theta_{max} = \frac{1}{n} \sum_{i=1}^n M_i \tag{10}$$

$$\theta_{std} = \frac{1}{m} \sum_{j=1}^m S_j \tag{11}$$

합성곱 신경망 모델은 신규 불량 패턴을 낮은 소프트 맥스 최댓값으로 예측해야 한다. 하지만 <Figure 10>에서 확인할 수 있듯이 합성곱 신경망 기반의 분류 모델의 특성상 특정 클래스에 대해 과신하는 경향이 있어 정확한 확률값을 나타내지 못하는 경우가 많아 신규 불량 패턴의 탐지 성능이 떨어진다.

하지만 본 연구에서는 사전에 정의된 불량 패턴만을 학습한

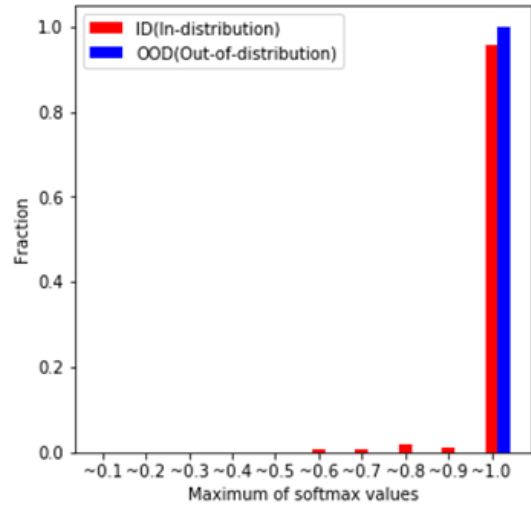
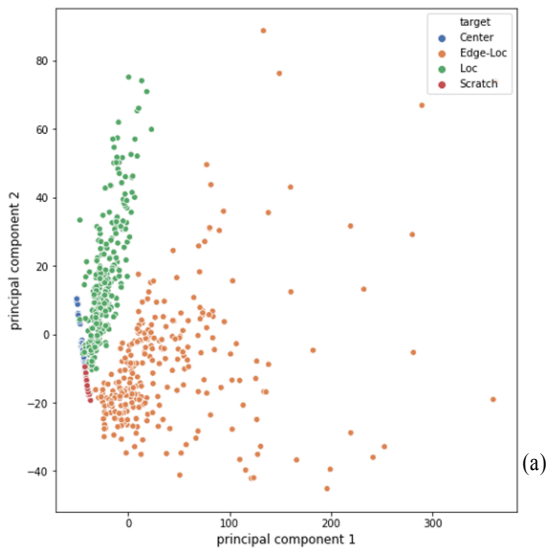
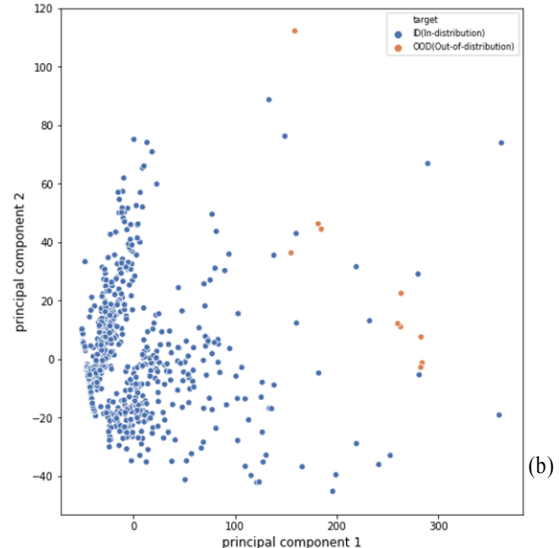


Figure 10. Distribution of Maximum Softmax Values

합성곱 신경망 모델을 통해 추출한 특징을 이용함으로써 분류 정확도와 신규 불량 패턴 탐지 성능을 향상하였다. 특징 추출 단계를 거침으로써 사전에 정의된 불량 패턴 중 같은 불량 패턴을 가지는 WBM끼리 유사한 특징을 얻을 수 있다. 반면 신규 불량 패턴이 특징 추출기를 통과할 경우에는 사전에 정의된 불량 패턴의 특징들과 구분되는 특징을 얻을 수 있다. <Figure 11>은 사전에 정의된 불량 패턴만을 학습에 이용한 합성곱 신경망 모델을 이용해 추출한 특징들을 principal component analysis(PCA)를 이용해 시각화한 그림 예시이다. <Figure 11(a)>를 통해 사전에 정의한 불량 패턴인 center, edge-loc, loc, scratch의 특징들이 같은 불량 패턴끼리 유사한 특징을 가지고 있음을 확인할 수 있다. 또한 <Figure 11(b)>를 통해 특징 추출기를 거친 신규 불량 패턴인 near-full, donut의 특징들은 사전에 정



Extracted features of predefined defect patterns



Extracted features of ID and OOD defect patterns

Figure 11. Extracted features of predefined defect patterns and OOD defect patterns using CNN

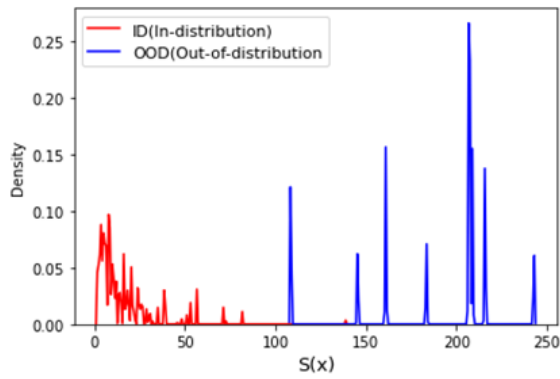


Figure 12. An Example Estimating $S(x)$ Distribution of ID and OOD

의되어 있는 불량 패턴의 특징과 구분되어 표현되어 있음을 확인할 수 있다. 각 불량 패턴의 구분되어 있는 특징을 multi-class data description 모델의 입력 샘플로 사용함으로써 기존 불량 패턴의 분류 및 신규 불량 패턴의 탐지 성능을 향상하였다.

또한 사전에 정의되어 있는 불량 패턴을 각 클래스별로 독립된 영역으로 형성하여 신규 불량 패턴의 발생을 가정함으로써 신규 불량 패턴의 발생을 가정하지 않는 분류 모델의 한계를 보완하였다. 형성된 영역을 바탕으로 각 샘플과 가장 가까운 클래스의 영역과의 거리 $S(x)$ 를 이용해 신규 불량 패턴을 탐지하였다. <Figure 12>는 사전에 정의한 불량 패턴과 신규 불량 패턴의 $S(x)$ 확률 분포를 KDE 방법을 이용해 추정한 예시이다. 사전

에 정의한 불량 패턴인 center, edge-loc, loc, scratch는 학습 데이터로 사용한 사전에 정의된 불량 패턴의 해당 클래스와 유사한 특징을 가지고 있기 때문에 해당 영역까지의 거리가 가까워 $S(x)$ 값의 대부분이 작은 값을 가지고 있다. 반면 신규 불량 패턴인 near-full, donut은 사전에 정의된 불량 패턴과 구분되는 특징을 가지고 있기 때문에 형성된 영역과 거리가 멀어 $S(x)$ 값이 대부분 큰 값을 가지고 있음을 확인할 수 있다. 사전에 정의된 불량 패턴과 신규 불량 패턴의 $S(x)$ 분포가 다르기 때문에 학습 데이터의 $S(x)$ 분포 상위 5%값을 임계값으로 설정하고, 임계값을 넘을 시 신규 불량 패턴으로 판단하였다.

최종적으로 <Table 4>와 같이 소프트웨어 확률을 기반으로 한 방법보다 본 연구에서 제안하는 방법이 기존 불량 패턴 분류 및 신규 불량 패턴 탐지에 있어 더 높은 성능을 보임을 확인할 수 있다.

4.5 클래스 수에 따른 불량 패턴 분류 및 탐지 성능 평가

제4.5절에서는 제안한 방법을 이용하여 edge-ring 패턴을 edge-loc 패턴으로 재분류하기 전과 후의 성능 비교와 사전에 정의한 불량 패턴의 수에 따른 성능 변화를 확인하기 위한 실험을 진행하였다. 실험은 4차에 걸쳐 진행하였으며 <Table 5>와 같이 사전에 정의한 불량 패턴의 수를 변경해가며 실험을 진행하였다. 1차 실험에서는 사전에 정의한 불량 패턴을 center, edge-loc, loc로, 신규 불량 패턴을 near-full, donut으로 가정

Table 4. Comparison of experimental results of proposed method and CNN based model

	Accuracy(%)	TP(%)	FN(%)	TN(%)	FP(%)
θ_{max}	85.2	80.4	19.6	0	100
θ_{std}	85.2	87.4	12.6	0	100
Proposed method	94.3	96.2	3.8	94.1	5.9

Table 5. Predefined and new defect patterns for the first, second, third and fourth experiments

	Predefined defect patterns	New defect patterns
1st Experiment	Center, Edge-Loc, Loc	Near-full, Donut
2nd Experiment	Center, Edge-Loc, Loc, Scratch	Near-full, Donut
3rd Experiment	Center, Edge-Loc(after reclassification), Loc, Scratch	Near-full, Donut
4th Experiment	Center, Edge-Loc, Edge-Ring, Loc, Scratch	Near-full, Donut

Table 6. Experimental Results of the First, Second, Third, and Fourth Experiments

	Accuracy(%)	TP(%)	FN(%)	TN(%)	FP(%)
1st Experiment	93.7	94.4	5.6	88.2	11.8
2nd Experiment	94.0	95.4	4.6	88.2	11.8
3rd Experiment	94.3	96.2	3.8	94.1	5.9
4th Experiment	91.1	94.9	5.1	94.1	5.9

하여 실험하였다. 2차 실험에서는 center, edge-loc, loc, scratch를 사전에 정의한 불량 패턴으로, near-full, donut을 신규 불량 패턴으로 가정하였다. 3차 실험에서는 center, edge-ring 패턴을 edge-loc 패턴으로 재분류한 후의 edge-loc, loc, scratch를 사전에 정의한 불량 패턴으로, near-full, donut을 신규 불량 패턴으로 가정하였다. 4차 실험에서는 center, edge-ring, edge-loc, loc, scratch를 사전에 정의한 불량 패턴으로, near-full, donut을 신규 불량 패턴으로 가정하여 실험을 진행하였다.

<Table 6>을 통하여 각 실험에 대한 결과를 확인할 수 있다. 3차, 4차 실험결과를 통해 edge-ring 패턴의 재분류 전과 후의 성능을 비교할 수 있다. edge-ring 패턴과 edge-loc 패턴의 일부 데이터가 비슷하기 때문에 재분류 전에 비해 사전에 정의한 불량 패턴의 분류 성능이 떨어지는 것을 확인할 수 있다. 또한 1차, 2차, 4차 실험 결과를 통해 사전에 정의한 불량패턴의 수에 따른 성능을 비교할 수 있다. 각 실험 결과를 통해 사전에 정의한 불량 패턴의 수에 따른 모델 성능의 변동성이 낮다는 것을 알 수 있다. 신규 불량 패턴은 기존 불량 패턴에 존재하지 않는 새로운 불량 유형의 특징을 가지고 있기 때문에 사전에 정의된 불량 패턴의 수에 적은 영향을 받음을 확인할 수 있다.

5. 결론

반도체 제조 공정 중 EDS 테스트의 결과로 얻어지는 WBM에 표시되는 각 칩의 불량 여부와 위치 정보를 이용해 불량 패턴을 판단할 수 있다. 같은 불량 패턴을 가지는 WBM은 특정 공정, 설비의 영향을 받았을 확률이 높으며 불량률의 원인 또한 동일한 가능성이 크다. 따라서 WBM의 불량 패턴을 정확하고 빠르게 분류하는 것은 중요하다. 대부분의 기존 연구들은 사전에 정의되어 있는 불량 패턴의 분류 성능을 높이는 방향으로 진행되어 왔기 때문에 신규 불량 패턴이 발생했을 시에 탐지하지 못한다는 한계가 있다. 신규 불량 패턴을 탐지하지 못했을 경우, 공정 이력을 잘못 추적하여 불량률의 원인을 잘못 판단할 수 있기 때문에 신규 불량 패턴을 탐지하는 것은 매우 중요하다. 따라서 본 연구에서는 사전에 정의된 불량 패턴만을 학습 과정에 이용하고, 학습 과정에서 각 불량 패턴별로 독립된 영역을 형성하여 신규 불량 패턴의 발생을 가정함으로써 사전에 정의된 불량 패턴을 분류함과 동시에 신규 불량 패턴을 탐지했다.

본 연구에서는 불량 패턴을 보다 명확하게 표현하기 위해 WBM의 노이즈로 판단되는 칩을 제거한 후, 합성곱 신경망 모델을 이용해 각 불량 패턴이 구분되어 표현될 수 있는 특징을 추출한다. 그 후 Deep-MCDD(Lee *et al.*, 2020) 기반의 multi-class data description 모델을 통해 사전에 정의된 불량 패턴의 특징을 각 불량 패턴별로 독립된 영역을 이루도록 학습하고, 각 샘플과 영역 사이의 거리를 이용해 사전에 정의된 불량 패턴을 분류함과 동시에 신규 불량 패턴을 탐지하였다. 실제

WBM 데이터를 이용해 평가한 결과, 우수한 성능으로 기존 불량 패턴을 분류하면서 신규 불량 패턴을 검출함을 확인하였다. 기존 불량 패턴 분류와 신규 불량 패턴 탐지가 가능한 모델을 이용해 불량 패턴 분류를 자동화할 수 있고, 오분류율을 줄임으로써 불량률의 원인 추적 및 개선 과정에 기여할 수 있다.

본 연구에서는 발생한 적이 있는 신규 불량 패턴에 대해서도 신규 불량 패턴으로 탐지한다는 한계가 있다. 하지만 불량 패턴이 추가될 때마다 전체 모델을 재학습하는 것을 비효율적이다. 따라서 신규 불량 패턴 데이터만의 재학습을 통해 발생한 적이 있는 신규 불량 패턴을 기존 불량 패턴으로 분류할 수 있는 모델로 확장해야 할 것이다.

참고문헌

- Baek, S. H., Lee, C. H., and Kim, S. B. (2020), Detection of a new pattern of wafer bin maps using deep learning, *Korean Journal of Industrial Engineering*, **46**(3), 326-337.
- Chen, F. L. and Liu, S. F. (2000), A neural-network Approach to Recognize Defect Spatial Pattern In Semiconductor Fabrication, *IEEE Transactions on Semiconductor Manufacturing*, **13**(3), 366-373.
- Higgins, J. H. (2004), *Introduction to Modern Nonparametric Statistics*, Belmont, CA, USA: Duxbury.
- Jang, J., Seo, M., and Kim, C. O. (2020), Support Weighted Ensemble Model for Open Set Recognition of Wafer Map Defects, *IEEE Transactions on Semiconductor Manufacturing*, **33**(4), 635-643.
- Jeong, Y. S., Kim, S. J., and 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.
- Jin, C. H., Na, H. J., Piao, M., Pok, G., and Ryu, K. H. (2019), A novel DBSCAN-based Defect Pattern Detection and Classification Framework for Wafer Bin Map, *IEEE Transactions on Semiconductor Manufacturing*, **32**(3), 286-292.
- Kang, S. (2020), Rotation-Invariant Wafer Map Pattern Classification with Convolutional Neural Networks, *IEEE Access*, **8**, 170650-170658.
- Kim, B., Jeong, Y. S., Tong, S. H., Chang, I. K., and Jeong, M. K. (2015), Step-down Spatial Randomness Test for Detecting Abnormalities in DRAM Wafers with Multiple Spatial Maps, *IEEE Transactions on Semiconductor Manufacturing*, **29**(1), 57-65.
- Kim, B., Jeong, Y. S., Tong, S. H., and Jeong, M. K. (2020), A Generalised Uncertain Decision Tree for Defect Classification of Multiple Wafer Maps, *International Journal of Production Research*, **58**(9), 2805-2821.
- Kim, Y., Cho, D., and Lee, J. H. (2021), Wafer Defect Pattern Classification with Detecting Out-of-Distribution, *Microelectronics Reliability*, **122**, 114157.
- Lee, D., Yu, S., and Yu, H. (2020), Multi-class Data Description for Out-of-Distribution Detection, In *Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, 1362-1370.
- Li, T. S. and Huang, C. L. (2009), Defect Spatial Pattern Recognition Using a Hybrid SOM-SVM Approach in Semiconductor Manu-

- facturing, *Expert systems with Applications*, **36**(1), 374-385.
- Liao, C. S., Hsieh, T. J., Huang, Y. S., and Chien, C. F. (2013), Similarity Searching for Defective Wafer Bin Maps in Semiconductor Manufacturing, *IEEE Transactions on Automation Science and Engineering*, **11**(3), 953-960.
- Liu, S. F., Chen, F. L., and Lu, W. B. (2002), Wafer Bin Map Recognition Using a Neural Network Approach, *International Journal of Production Research*, **40**(10), 2207-2223.
- Nakazawa, T. and Kulkarni, D. V. (2018), Wafer Map Defect Pattern Classification and Image Retrieval Using Convolutional Neural Network, *IEEE Transactions on Semiconductor Manufacturing*, **31**(2), 309-314.
- Piao, M., Jin, C. H., Lee, J. Y., and Byun, J. Y. (2018), Decision Tree Ensemble-Based Wafer Map Failure Pattern Recognition Based on Radon Transform-Based Features, *IEEE Transactions on Semiconductor Manufacturing*, **31**(2), 250-257.
- Saqlain, M., Jargalsaikhan, B., and Lee, J. Y. (2019), A Voting Ensemble Classifier for Wafer Map Defect Patterns Identification in Semiconductor Manufacturing, *IEEE Transactions on Semiconductor Manufacturing*, **32**(2), 171-182.
- Saqlain, M., Abbas, Q., and Lee, J. Y. (2020), A Deep Convolutional Neural Network for Wafer Defect Identification on an Imbalanced Dataset in Semiconductor Manufacturing Processes, *IEEE Transactions on Semiconductor Manufacturing*, **33**(3), 436-444.
- Wang, J., Yang, Z., Zhang, J., Zhang, Q., and Chien, W. T. K. (2019), AdaBalGAN: An Improved Generative Adversarial Network with Imbalanced Learning for Wafer Defective Pattern Recognition, *IEEE Transactions on Semiconductor Manufacturing*, **32**(3), 310-319.
- Yoo Y., Ahn D., Park S. H., and Baek J. G. (2015), Wafer Map Image Analysis Methodology in Semiconductor Process, *Journal of Industrial Engineering*, **41**(3), 267-274.
- Yuan, T., Kuo, W., and Bae, S. J. (2011), Detection of Spatial Defect Patterns Generated in Semiconductor Fabrication Processes, *IEEE Transactions on Semiconductor Manufacturing*, **24**(3), 392-403.

저자소개

송창용 : 한양대학교 산업경영공학과에서 학사학위를 취득하고, 석사과정에 재학 중이다. 연구 분야는 머신러닝, 데이터마이닝이다.

정영선 : 려거스 대학교에서 산업시스템공학 박사학위를 취득하고, 현재 전남대학교 산업공학과에 재직 중이며 주요 관심 분야는 반도체 제조 공정의 통계적 데이터 마이닝 모델 개발, 지능형 운송 시스템이다.

김병훈 : 려거스 대학교에서 산업시스템공학 박사학위를 취득하고, 현재 한양대학교 산업경영공학과에 재직 중이며 주요 관심 분야는 통계 데이터 마이닝 방법론 개발, 반도체 제조 공정용 데이터 마이닝 모델 개발, 그래프 마이닝이다.