

반도체 설비 센서 데이터를 활용한 딥러닝 기반의 불량예측 모델에 관한 연구

하승재*, 이원석*, 구교연*, 신용태**
*승실대학교 IT정책경영학과

sjha0214@naver.com, wsmjmy@naver.com, shin@ssu.ac.krr

A Study on the Deep Learning-Based Defect Prediction Model Using Sensor Data of Semiconductor Equipment

Seung-Jae Ha*, Won-Suk Lee*, Kyo-Yeon Gu*, Yong-Tae Shin**

*Dept. of IT Policy and Management, Graduate School of Soongsil University

**Dept. of Computer Science, Soongsil University

요 약

본 연구는 반도체 제조 공정중 발생하는 센서 데이터를 활용하여 딥러닝기반으로 불량을 예측하는 모델을 제안한다. 반도체 공장에서는 FDC(Fault Detection and Classification)라는 불량을 예측하는 시스템이 있지만, 공정의 복잡도가 높고 센서의 종류가 많아 공정 관리자가 모든 센서의 기준을 설정 및 관리하는데 한계가 있다. 이를 해결하기 위해 공정 설비의 센서 데이터를 딥러닝을 활용하여 학습시켜 센서 기준정보로 임계치를 제공하고, 가공중 발생하는 센서 데이터가 입력되면 정상 여부를 판정하는 모델을 제안한다.

1. 서론

반도체 제조업체들은 매우 경쟁이 치열한 글로벌 환경에서 수익성을 유지하기 위해 높은 수율 유지, 불량품 처리 및 역동적인 제조 환경을 고려한 프로세스 개선 등 수많은 문제를 해결해야 한다[1]. 본 연구는 반도체 공정에서 발생하는 센서 데이터를 활용하여 불량을 예측하고자 하는 것이다. 반도체 공장에서 불량을 예측하는 것은 품질관리 및 생산수율과 직결되는 매우 중요한 문제이다. 또한, 반도체공장에서 발생하는 시계열 데이터는 레이블링이 어렵고[2], 정상대비 이상의 숫자가 매우 적어 데이터의 불균형이 심한 특성이 있다. 기존 연구에서도 AI를 활용한 불량 예측에 관한 연구를 많이 진행하였지만, 대부분 불량원인에 대한 설명이 부족한 부분이 있어서[2], 이를 해결할 수 있는 방안을 알아보하고자 본 연구를 진행하였다.

2. 반도체공장 불량 예측모델 요구사항

반도체공장의 공정에서 발생하는 데이터의 특징인 첫째로 시간의 길이가 일정하지 않은 시계열 데이터, 둘째로 데이터의 레이블링이 어려운 점, 셋째로 정상 데이터대비 이상 데이터의 숫자가 지극히

적은 점, 즉, 데이터의 불균형이 심한 특성이 존재한다. 이러한 반도체 공장의 데이터 특성을 고려하여, 첫째, 시간의 길이가 일정하지 않은 시계열 데이터를 처리할 수 있고, 둘째, 레이블링되지 않은 데이터의 처리가 가능하고, 셋째, 정상과 이상의 데이터의 심한 불균형을 처리할 수 있고, 넷째, 불량원인을 추적 관리가 가능해야 한다는 것을 본 연구의 요구사항으로 전제하였다.

3. 선행 연구

3.1 Topology 및 Deep SVDD를 활용한 불량예측

본 연구를 진행함에 있어서 제일 먼저 시도 했던 방식으로 데이터의 전처리는 Topology(위상학적인 특징)를 활용하고 Deep SVDD를 이용하여 불량을 예측하는 연구를 진행하였다. TDA(Topology Data Analysis)는 Data가 갖고 있는 Topology를 추출하여 Data를 추출하는 기법으로 Raw Data에 sampling noise, white noise에 강건하다는 장점이 있다[3]. 또한 데이터의 학습은 범용적인 One-class anomaly detection 알고리즘으로 Deep SVDD를 활용하였다. 영구결과로 Persistence Topology를 구성하는 demension=3, skip=1, delay=1인 데이터 셋을 이용한

결과, 최적의 파라미터 $\nu=0.3$ 에서의 성능은 [표 1]과 같다.

[표 1] Topology 및 Deep SVDD 연구 결과

구분		예측		딥러닝 학습시간
		정상	불량	
실제	정상	536	1,377	1 hr (3,600 sec)
	불량	2	6	

연구 결과를 보면 Topology를 활용한 방식이 노이즈에 강한 특성이 있지만 정상과 불량 데이터 간의 차이가 작을 경우 판별의 어려움이 있어 정상 판정이 제대로 되지 않음을 확인되었다. 또한, 정상과 이상의 판정기준을 위한 딥러닝 학습시간이 약 1시간으로 너무 길어 학습시간 단축이 해결과제이었다.

3.2 FFT 및 Deep SVDD를 활용한 불량예측

앞의 선행 연구에서 데이터 간의 차이가 작을 경우를 해결하기 위해 데이터 전처리 방식으로 FFT(Fast Fourier Transfor, 고속 푸리에 변환)를 선택하였다. FFT는 시계열 데이터에서 보이지 않던 부분을 보이게 하는 장점이 있어 본 연구에 활용하였다[4]. FFT로 전처리한 데이터를 Deep SVDD로 학습시켜 불량을 예측하였으며[5], 그 연구결과는 [표 2]와 같다.

[표 2] FFT 및 Deep SVDD 연구 결과

구분		예측		딥러닝 학습시간
		정상	불량	
실제	정상	1,676	715	10 min (600 sec)
	불량	2	5	

[표 2]를 보면 Topology 및 Deep SVDD를 활용한 방식에 비해 성능이 많이 개선되었으며, 학습시간도 10분으로 많이 개선되었다. 하지만 Deep SVDD의 특성상 불량의 원인을 분석하고자 할 경우에 추적관리의 어려움이 존재하여 원인분석을 가능하도록 하는 것이 해결과제이었다.

3.3 LSTM 기반 Autoencoder를 활용한 불량예측

다음으로 진행한 연구는 앞의 연구에서 문제이었던 불량 발생원인 분석이 불가능한 부분을 해결하기 위하여 불량원인 추적이 가능한 LSTM 기반의 Autoencoder를 활용하였다[6]. 반도체 제조공정의 1회 사이클의 센서 데이터를 분석하여 공정중에서 센서값이 급격하게 변화하는 시점을 기준으로 공정진, 공정중, 공정후로 구분하여 Step 1 Step 2, Step 3로

데이터를 세분화하였다. 또한, Step 1, Step 3은 노이즈가 심하게 생기는 경우가 많아서 Step 2로 학습 및 평가를 진행하였다. 본 연구에서 활용한 LSTM 기반 Autoencoder는 Sequence Data에 Encoder-Decoder LSTM 아키텍처를 적용하여 구현한 Autoencoder이다. 연구를 진행하면서 Raw 데이터 값이 매우 작은 값으로 gradient가 거의 0에 가까운 값이어서 학습이 잘 진행되지 않아 원래 Raw 데이터에 100을 곱한 값을 이용하여 학습을 진행하였다. 적절한 학습시간을 고려할 경우 3 Layer의 LSTM Autoencoder의 파라미터는 256차원을 가진 Hidden Layer가 가장 좋은 성능을 나타냈으며, 테스트 데이터로 최적의 threshold를 설정하여 분류한 연구결과는 [표 3]과 같다.

[표 3] LSTM 기반 Autoencoder 연구 결과

구분		예측		딥러닝 학습시간
		정상	불량	
실제	정상	2,815	53	5 min (300 sec)
	불량	2	5	

LSTM 기반 Autoencoder의 연구결과 앞의 두 방법에 비해 불량원인 추적이 가능하고, 성능이 우수하였다. 하지만 학습속도가 5분으로 1개 라인에 대한 학습시간을 계산하며 약 8일(17개 센서, 학습시간 5분, 레시피 종류 15개, 챔버수 3개, Step 3 가정시)이 소요되어 실제 제조현장에 적용함에 있어서 어려움이 존재한다.

4. 데이터 소개 및 불량예측 모델

4.1 데이터 수집과 전처리

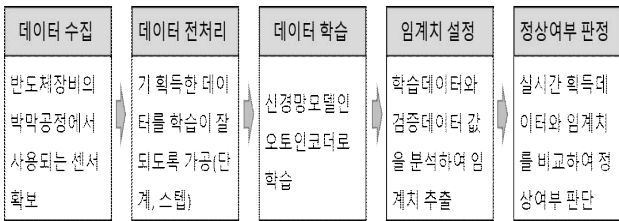
반도체 박막공정의 제조설비에서 발생하는 압력, 스피드, 가스, 전압, 필름두께 등 17개 유형의 센서 데이터를 입력데이터로 28,678개를 수집하였고, 테스트 데이터는 7개로 그중 정상 데이터 2개, 불량 데이터 5개였다. 또한, 데이터의 구조는 [표 4]와 같다.

[표 4] 데이터 구조

Field	Description
Equip	장비명
Date & Time	이력 발생시간
Setting Chamber	세팅 제품라인
Actual Chamber	실제 제품라인
Recipe	레시피명
Lot	Lot 번호
Wafer	웨이퍼 번호
Value	데이터값

반도체 박막공정의 센서 데이터는 동일한 레시피라도 데이터가 동일하지 않고, 데이터에도 Sampling noise, white noise, random noise 등의 다양한 노이즈가 존재하며, 정상 데이터 대비 비정상 데이터가 매우 적어 데이터의 불균형이 존재하는 특징이 있다. 이러한 데이터의 특성을 고려하여 제조공정에서 수집된 데이터는 분석을 위하여 다음과 같은 전처리를 수행하였다. 첫째, 수집된 데이터중 설비 및 레시피별 자료를 분류하여 동일한 설비 및 레시피에 대해 실험을 진행하였다. 둘째, 제조공정의 센서데이터는 시계열 데이터로 제조시마다 발생하는 시간 차이를 표준화하여 분석을 진행하였다.

4.2 연구 진행순서



[그림 1] 반도체 불량 예측모델 연구 진행순서

본 연구에서 수행한 딥러닝 기반의 반도체 제조공정의 불량예측 모델의 연구 진행순서는 [그림 1]과 같다. 그 진행순서는 반도체 제조설비의 제조공정에서 사용되는 센서로부터 기 획득된 학습데이터를 준비하는 데이터 수집단계와, 수집된 데이터를 학습이 잘 되도록 수집 데이터의 구조 및 값을 변환하는 전처리 단계, 전처리한 데이터를 뉴럴 네트워크 모델 기반의 오토인코더로 학습시키는 데이터 학습 단계와, 학습데이터와 오토인코더로 학습한 결과데이터를 비교하여 학습결과를 확인 및 학습데이터와 결과데이터의 값을 분석하여 임계치를 추출하는 임계치 설정 단계 및 제조공정중 센서로부터 실시간 획득되는 센서값이 임계치 이상일 경우 비정상 공정으로 판단하고, 임계치 미만일 경우 정상 공정으로 판정하는 단계를 거친다.

정 단계 및 제조공정중 센서로부터 실시간 획득되는 센서값이 임계치 이상일 경우 비정상 공정으로 판단하고, 임계치 미만일 경우 정상 공정으로 판정하는 단계를 거친다.

4.3 불량예측모델

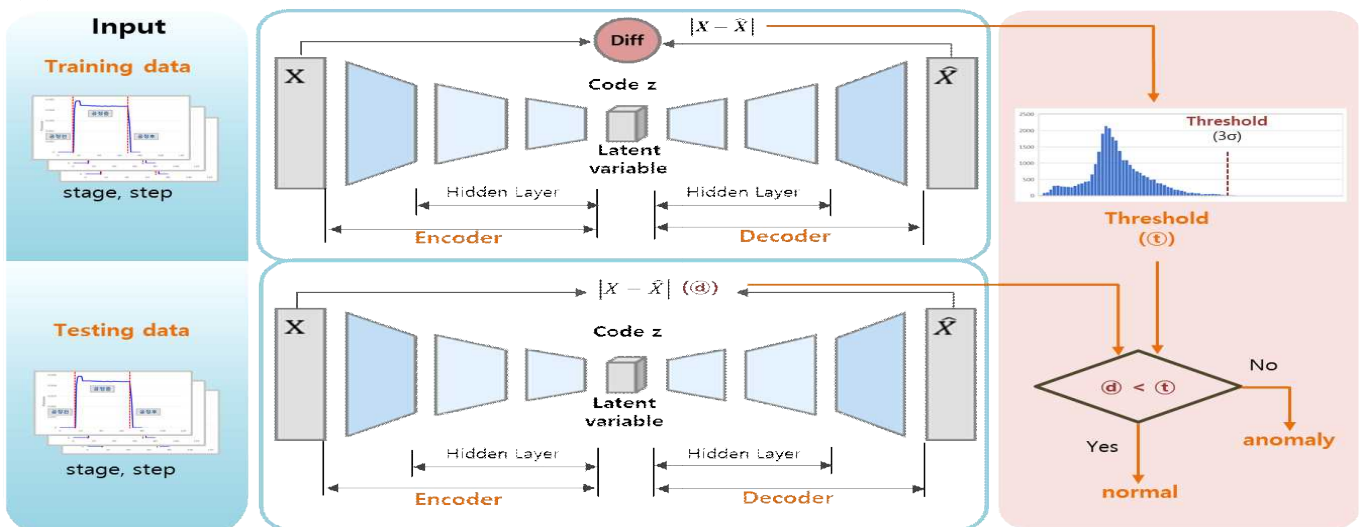
본 연구의 불량예측 모델이란 제조공정 과정에서 발생하는 설비의 센서 데이터를 딥러닝 기법으로 분석하여 제품의 이상을 예측하는 모델로 정의하며, 모델의 이미지는 [그림 2]와 같다.

5. 연구결과

선행 연구에서 진행한 Topology & Deep SVDD, FFT & Deep SVDD, LSTM based Autoencoder, 반도체 장비인 FDC, 본 연구의 불량예측 모델에 대해 분류평가표에 의한 성능 평가, 학습속도 및 불량 원인분석 가능여부에 대해 종합적으로 정리하면 [표 5]와 같다.

[표 5] 불량예측 모델 연구결과

구분	Accuracy	Precision	Recall	F1-score	학습속도	원인분석
Topology & Deep SVDD	0.2821	0.9963	0.2802	0.4374	1 hour (3,600 sec)	X
FFT & Deep SVDD	0.7009	0.9988	0.7009	0.8237	10 min (600 sec)	X
LSTM based Autoencoder	0.9809	0.9993	0.9815	0.9903	5 min (300 sec)	O
FDC (반도체 장비)	0.9794	1.0000	0.9794	0.9896	N/A	O
불량예측 모델	0.9994	0.9999	0.9994	0.9997	15 sec	O



[그림 2] 반도체 설비 센서 데이터를 활용한 딥러닝 기반의 불량예측 모델

연구 결과를 보면 본 연구의 불량예측 모델이 FDC의 Precision이외에는 모든 지표에서 탁월하게 우수한 결과를 나타내고 있다. 특히, 반도체 장비인 FDC에 비해서도 동일 또는 우수한 성능을 나타내고 있어서 반도체 공정에 현장적용이 가능한 수준이고, 학습속도는 15초수준으로 1개 라인에 대해 학습시간이 약 8시간내외로 1일이면 학습을 완료할 수 있어 학습속도 또한 현장 적용이 가능한 수준이다. 또한, 불량 원인분석도 가능하여 본 연구의 불량예측 모델이 모든 면에서 우수한 것으로 나타났다.

6. 결론

본 연구는 반도체 박막 제조공정에서 사용되는 센서 데이터를 활용하여 딥러닝 기반의 최적의 불량예측모델을 개발하고자 하였다. 선행 연구로 Topology & Deep SVDD, FFT & Deep SVDD, LSTM based Autoencoder를 활용한 학습모델을 검증하였다. 이를 기반으로 부족한 부분을 개선하여 본 연구의 불량예측 모델을 완성하였다. 선행연구 모델 3가지와 과 본 연구모델 및 반도체 장비인 FDC와 성능을 비교분석 하였다. 그 결과 분류평가표에 의한 성능 평가, 학습속도 및 불량 원인분석 가능여부에 대해 종합적으로 정리하면 본 연구의 불량예측 모델이 모델 성능지표인 Accuracy, Precision, Recall, F1-Score 모두 0.999 수준으로 탁월하게 우수한 결과를 나타냈다. 특히, 반도체 장비인 FDC에 비해서도 동일 또는 우수한 성능을 나타내어 반도체 공정에 현장적용이 가능한 수준이다. 또한, 불량 원인분석도 가능하여 본 연구의 불량예측 모델이 모든 면에서 우수한 것으로 나타났다. 하지만 본 연구의 실제 결과 데이터 7개로 연구 결과를 정확히 평가함에 있어서는 한계가 존재한다. 이에 대해서는 향후 추가로 검증할 필요가 있다고 판단된다.

참고문헌

- [1] Moyne, J., & Iskandar, J. (2017). Big data analytics for smart manufacturing: Case studies in semiconductor manufacturing. *Processes*, 5(3), 39.
- [2] Ruff, L., Kauffmann, J. R., Vandermeulen, R. A., Montavon, G., Samek, W., Kloft, M., Müller, K.-R. (2021). A unifying review of deep and shallow anomaly detection. *Proceedings of the IEEE*.
- [3] Umeda, Y. (2017). Time series classification via topological data analysis. *Information and Media Technologies*, 12, 228-239.
- [4] Ruff, L., Vandermeulen, R., Goernitz, N., Deecke, L., Siddiqui, S. A., Binder, A., Kloft, M. (2018). Deep one-class classification. Paper presented at the International conference on machine learning.
- [5] Ruff, L., Vandermeulen, R. A., Görnitz, N., Binder, A., Müller, E., Müller, K.-R., & Kloft, M. (2019). Deep semi-supervised anomaly detection. *arXiv preprint arXiv:1906.02694*.
- [6] Malhotra, P., Ramakrishnan, A., Anand, G., Vig, L., Agarwal, P., & Shroff, G. (2016). LSTM-based encoder-decoder for multi-sensor anomaly detection. *arXiv preprint arXiv:1607.00148*.