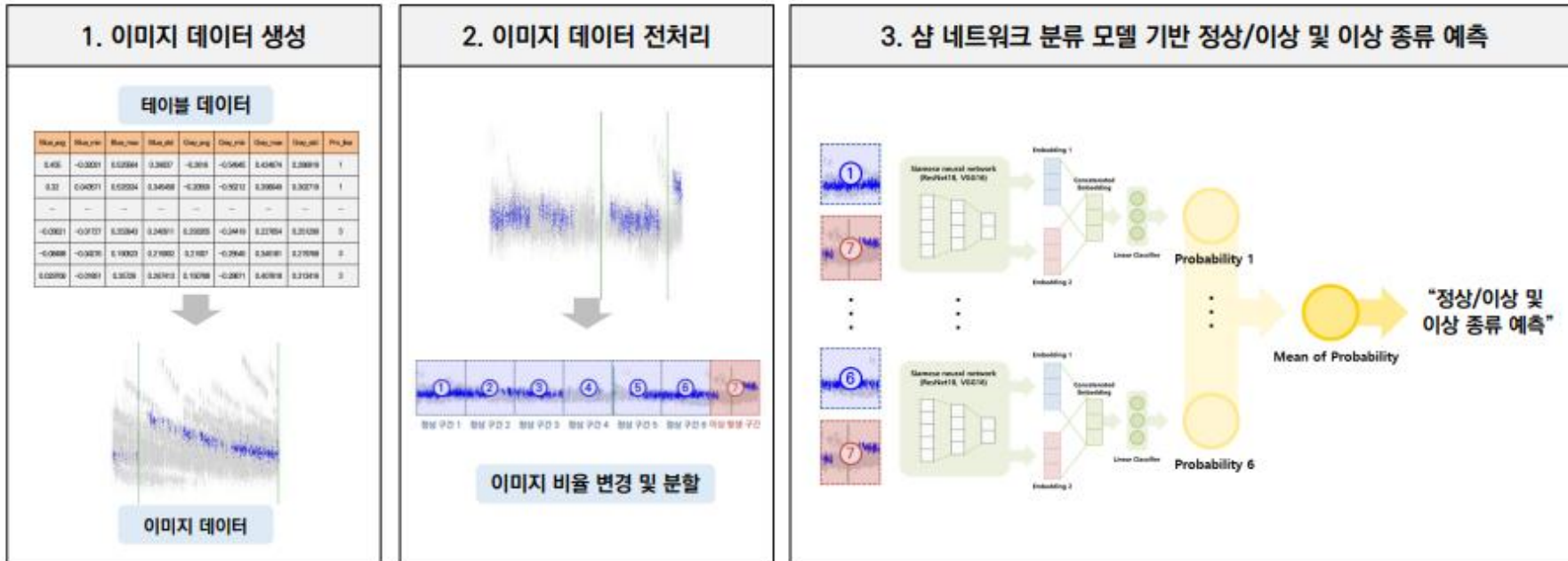


# 산공학회건문록

# 이상치 탐지

## 삼 네트워크 모델 기반 이상탐지

- 테이블 데이터를 이미지 데이터로 변경
- 특정 구간 별로 시계열 이미지를 구간 나눔 ※이때 나누는 텀은 사용자의 정의에 따라 세팅
- 삼 네트워크 모델을 활용해 특정 구간이 이상 및 이상 종류 예측 수행 ※이때 앵커 이미지는 사용자 재량



# 이상치 탐지

## 삼 네트워크 모델 기반 이상탐지

[ 분류 성능 ]

VGG16

Accuracy	Macro F1 score	Recall
0.8046	0.8047	0.8582

※ 당해년도 목표치 : 80%

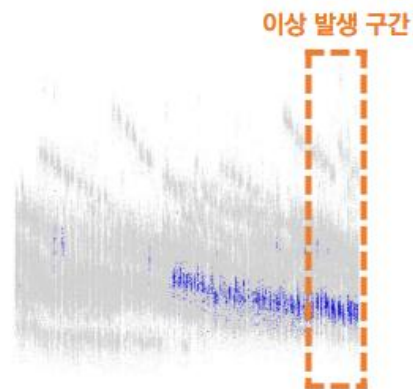
[ Confusion Matrix ]

VGG16

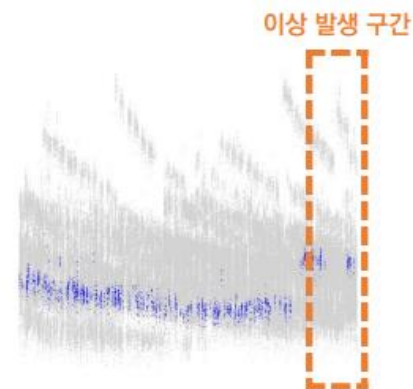
		예측 값	
		정상	이상
실제 값	정상	92	32
	이상	18	109

- ✓ 정상으로 잘 분류된 정상
  - 앞 구간의 패턴들과 유사한 패턴을 보임
- ✓ 이상으로 잘못 분류된 정상
  - 앞 구간의 패턴들과 다른 패턴을 보임

[ 이상으로 잘 분류된 DEV ]



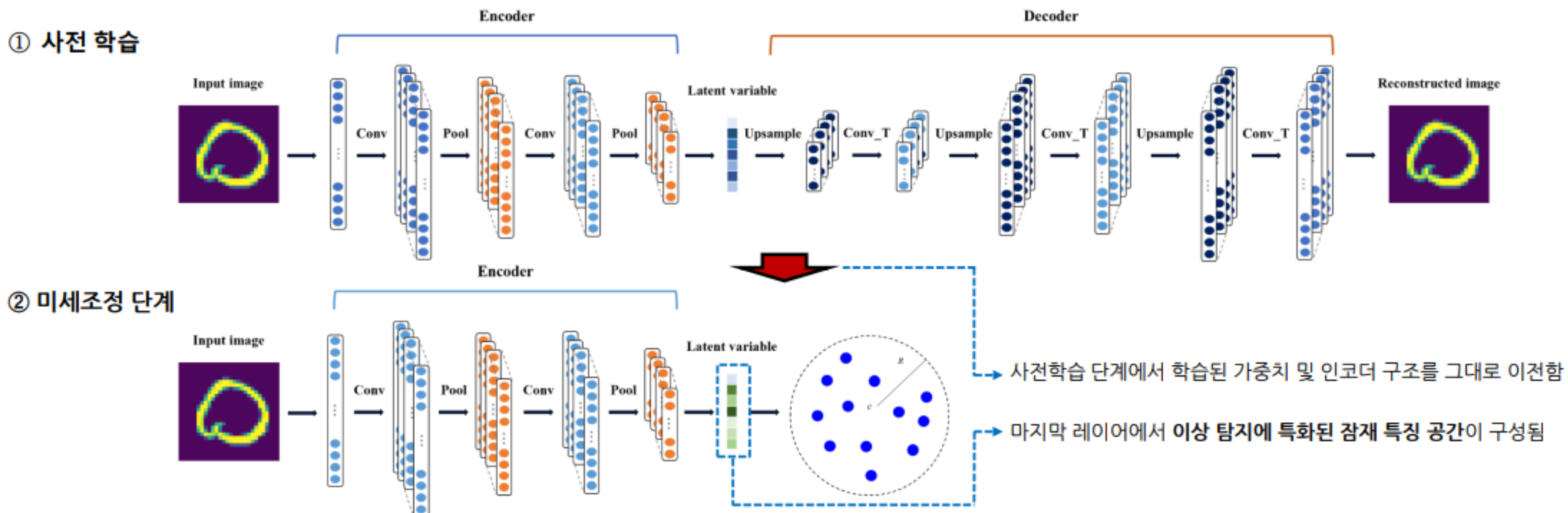
[ 정상으로 잘못 분류된 DEV ]



- 여러 시계열 센서 데이터를 요약할 수 있는 scatter plot 형태의 이미지로 변환 => 과거 시점과 최근 시점 간의 차이를 비교할 수 있도록 전처리 가능!
- 삼 네트워크 구조로 간단하면서 더 좋은 성능을 보이는 모델 구축 => semi/self – supervised learning 또한 도입 가능!

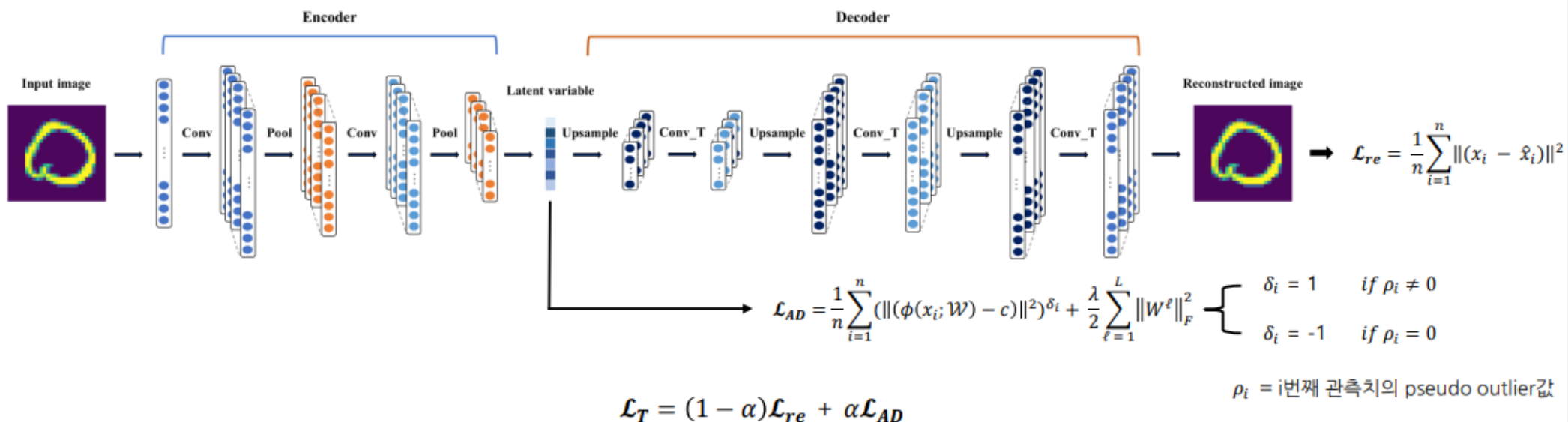
## Deep One-class Classification with Pseudo Outliers for Robust Anomaly Detection

- 기존 SVDD의 fine tuning 단계에서의 loss는 잡음 관측치의 영향을 받아 왜곡된 구체를 생성할 수 있음
- 이를 해결하기 위해 pseudo outlier 기반의 deep SVDD 방법론 제시



## Deep One-class Classification with Pseudo Outliers for Robust Anomaly Detection

- 기존 오토인코더에 주로 활용되는 재구축 오차 ( $\mathcal{L}_{re}$ )와 기존 deep SVDD의 손실 함수를 변형한 손실 함수( $\mathcal{L}_{AD}$ )를 통해 손실 함수( $\mathcal{L}_T$ )를 정의
- $\mathcal{L}_T$ 를 통해 도출되는 이상 탐지에 특화된 잠재 특징공간에서 초구체의 중심  $c$ 를 계산하고, 학습 과정에 따라 갱신함 (초기 초구체의 중심  $c$ 는 랜덤값으로 정의함)
- 단계별 학습 (사전 학습 + 미세 조정)없이도 단일 학습 과정을 통해 우수한 이상 탐지가 가능한 결정 경계선을 도출 가능



## Deep One-class Classification with Pseudo Outliers for Robust Anomaly Detection

- 잡음 데이터에 대한 강건성을 보다 명확하게 보여주기 위해 Deep SVDD의 손실 함수를 변형하여 제안 방법론의 손실 함수를 정의함

Deep SVDD의 손실 함수

$$\mathcal{W}^* = \underset{\mathcal{W}}{argmin} \frac{1}{n} \sum_{i=1}^n (\|\phi(x_i; \mathcal{W}) - c\|^2) + \frac{\lambda}{2} \sum_{\ell=1}^L \|\mathcal{W}^\ell\|_F^2$$



제안 방법론의 손실 함수

$$\mathcal{W}^* = \underset{\mathcal{W}}{argmin} (1 - \alpha) \frac{1}{n} \sum_{i=1}^n \|x_i - \hat{x}_i\|^2 + \alpha \left\{ \frac{1}{n} \sum_{i=1}^n (\|\phi(x_i; \mathcal{W}) - c\|^2)^{\delta_i} + \frac{\lambda}{2} \sum_{\ell=1}^L \|\mathcal{W}^\ell\|_F^2 \right\} \begin{cases} \delta_i = 1 & \text{if } \rho_i \neq 0 \\ \delta_i = -1 & \text{if } \rho_i = 0 \end{cases}$$

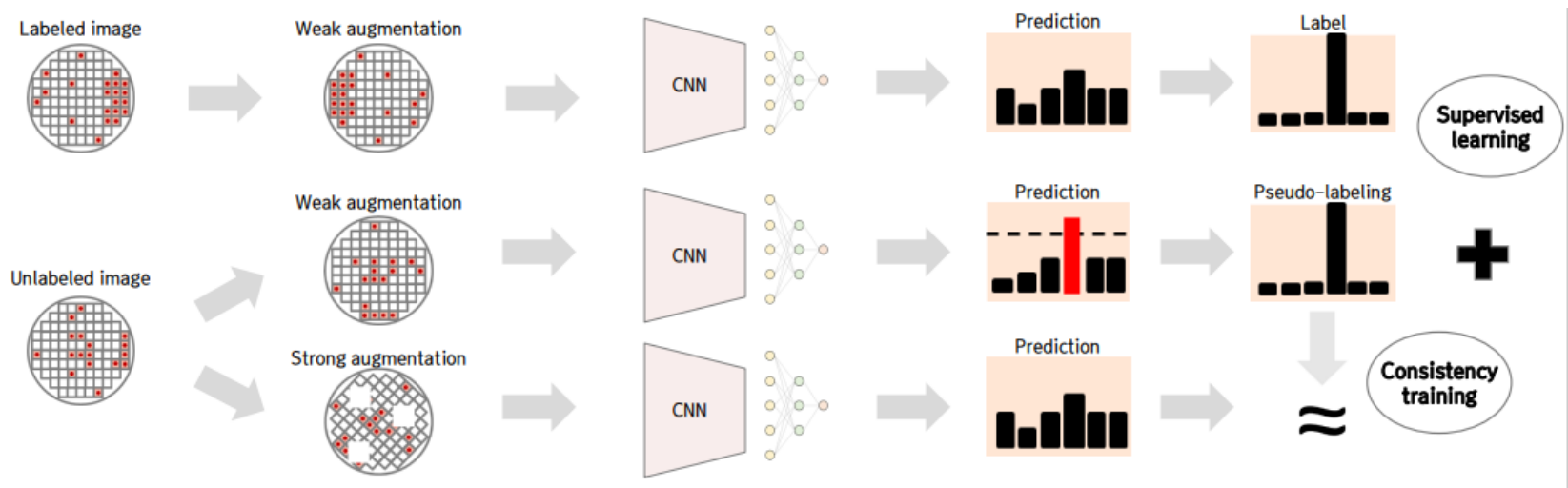
- 
- 잠재 특징공간에서 도출된 Pseudo outlier를 통해 각 데이터가 잡음 데이터에 가까운 데이터인지 아닌지를 판단하여 이를 손실 함수에 적용함
    - Pseudo outlier가 0이 아닌 데이터는  $\delta_i = 1$ 을 부여 받아 오차 함수를 최소화 하기 위해 **초구체의 중심에 더 가까워지도록 매핑**
    - Pseudo outlier가 0인 데이터는  $\delta_i = -1$ 을 부여 받아 오차 함수를 최소화 하기 위해 **초구체의 중심에서 더 멀어지도록 매핑**
  - 이상 탐지에 특화된 손실 함수와 재구축 손실 함수를 동시에 활용함으로써 **각 손실함수에서 생성되는 특징이 동시에 포함된 잠재 특징 공간을 구축 가능**

제안 방법론의 구조를 통해 도출된 잠재 특징공간에서의 상호 k-인접 이웃의 개수로 각 관측치의 Pseudo outlier를 정의

## 범주형 불균형이 반영된 준지도 학습을 활용한 반도체 웨이퍼 불량 패턴 검출

- 준지도학습(FixMatch[Google Research, 2020])를 활용해 wafer bin map 분류 성능 향상이 있었음
- FixMatch: pseudo-labeling과 consistency loss(unlabeled data의 raw data와 augmentation한 데이터의 softmax probability값의 차이를 줄여주는 방법으로 학습 진행)
- Consistency training을 하기 위해선 데이터에 맞는 augmentation 방법을 찾는 것이 중요

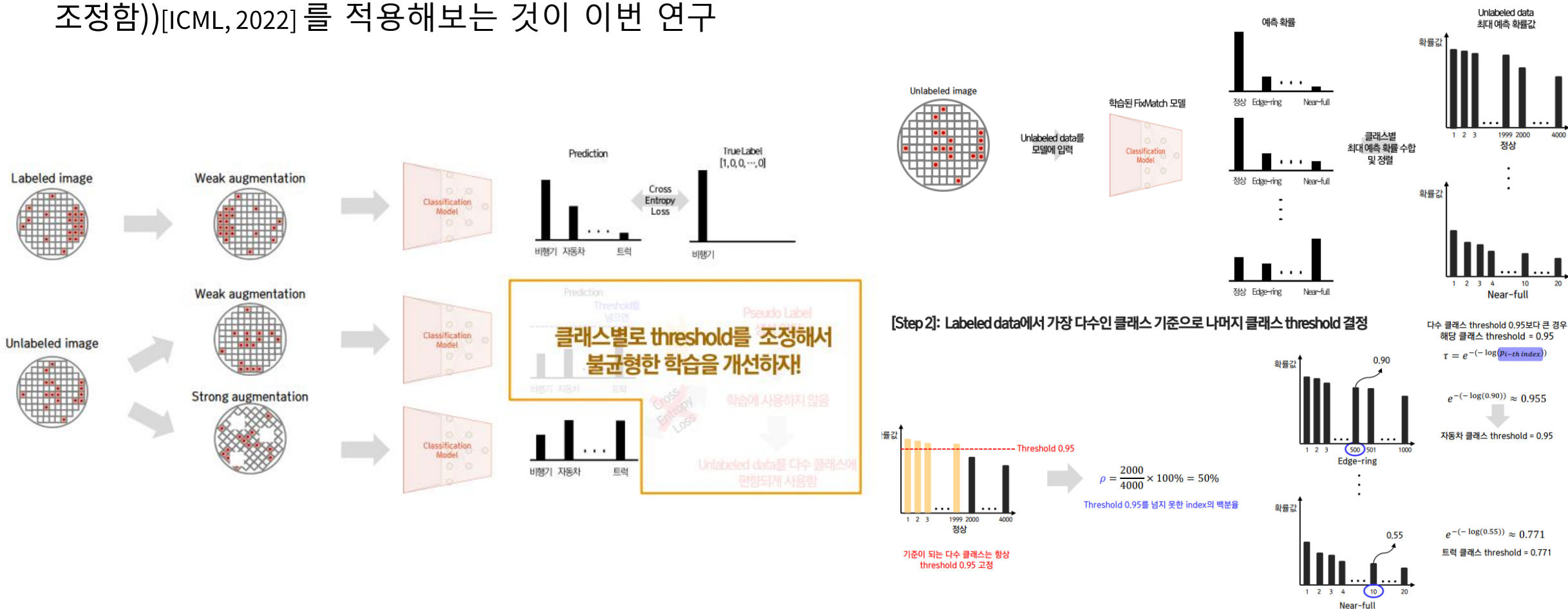
※ Consistency training : 다양한 형태의 데이터 변형을 학습하도록 하는 것





## 범주형 불균형이 반영된 준지도 학습을 활용한 반도체 웨이퍼 불량 패턴 검출

- Adsh(FixMatch + class adaptive thresholding(클래스별로 불균형 상황을 반영하여 threshold를 유연하게 조정함))[ICML, 2022] 를 적용해보는 것이 이번 연구





# 이상치 탐지

## 범주형 불균형이 반영된 준지도 학습을 활용한 반도체 웨이퍼 불량 패턴 검출

- FixMatch augmentation 조합에선 Rotation + Cutout&noise 조합이 가장 성능이 좋았음
- FixMatch 와 Adsh 비교 결과, 불균형 상황(1% 데이터만 사용)이 반영되었지만 성능 개선이 보이지 않았다
- EWMA와 같은 제약식 및 하이퍼파라미터 튜닝 시 결과 85% 정도로 성능 향상

## 실험 결과

가설 1: Unlabeled 데이터를 활용하여 준지도 학습을 한다면 지도 학습보다 효과적이고 데이터 부족 문제 해결 가능(가설 2)

❖ FixMatch augmentation 성능 비교

- Random select [Crop, shift, cutout, noise, shear] 목록에서 랜덤하게 추출 후 적용

Encoder Architecture	Training Method	Weak Augmentation	Strong Augmentation	Labeled Data			Test F1 score
				1,30,01%	6,19,02%	13,08,02%	
ResNet18	FixMatch	Rotation	Cutout + Noise	0.571	0.601	0.716	0.779
			Random select	0.603	0.625	0.730	0.808
			Random select (no crop, shear)	0.608	0.734	0.749	0.808
			Cutout + Noise	0.641	0.773	0.794	0.825
			Random select	0.559	0.658	0.720	0.771
		Flip	Random select (no crop, shear)	0.629	0.713	0.713	0.782

**Threshold 안정화를 위해서 Adsh에 제약식 적용 실험 결과**  
가설 2: 분포 보정에 활용할 수 있을 것임(가설 1)은 분포 보정 가능(가설 2)이다

Adsh에 제약식 적용 비교 실험 결과

- Exponentially weighted moving average를 사용하여 threshold smoothing을 하겠다 경우 좋은 성능 향상을 보임
- EWMA는  $\text{EWMA} = \text{EWMA} \times \text{threshold} + (1 - \text{EWMA}) \times \text{new\_threshold}$  형태 사용 → Exponentially weighted moving average

Training Method	Weak Augmentation	Strong Augmentation	Labeled Data			Test F1 score
			1,30,01%	6,19,02%	13,08,02%	
FixMatch	Rotation	Cutout + Noise	0.702 (0.600)	0.791 (0.600)	0.836 (0.600)	0.804 (0.600)
Adsh	Rotation	Cutout + Noise	0.699 (0.600)	0.790 (0.600)	0.835 (0.600)	0.803 (0.600)
Adsh + EWMA 0.5	Rotation	Cutout + Noise	0.715 (0.600)	0.819 (0.600)	0.858 (0.600)	0.863 (0.600)

Training Method	Weak Augmentation	Strong Augmentation	Labeled Data	Labeled Data	Labeled Data	Test F1 score
Adsh + EWMA 0.1	Rotation	Cutout + Noise	0.711	0.826	0.861	0.867
Adsh + EWMA 0.2	Rotation	Cutout + Noise	0.714	0.827	0.873	0.864
Adsh + EWMA 0.8	Rotation	Cutout + Noise	0.713	0.806	0.848	0.870
Adsh + EWMA 0.9	Rotation	Cutout + Noise	0.709	0.825	0.840	0.861

**결론**

- WM-811k 데이터셋에 불균형 상황이 반영된 준지도 학습을 적용
- FixMatch 준지도 학습 방법론을 적용하여 기존 지도 학습 대비 효과적인 분포 보정 성능 향상을 보임
- 기존 FixMatch에 class-imbalance를 반영하여 adaptive threshold를 적용하였을 경우, 더 우수한 분포 보정 성능 향상을 보임

## 시계열 기반 이상치 탐지 & 데이터셋 부족 문제

- 샘플 별 대조 학습 기법 기반 과대표집을 활용한 설비 상태 진단 프레임워크  
: MOCO[Kaiming He, 2020] 를 활용해서 베어링 회전 데이터셋에 적용했더니 다른 오버샘플링 기법보다 성능이 좋았다 => 시계열 기반의 설비 신호 데이터는 MOCO가 더욱 효과적이다
- Contrastive learning for time series anomaly detection using decomposition methods  
: 시계열 데이터를 분해한 뒤 분해된 시계열을 증강 후 대조학습을 진행하여 기존 신호와 다른 부분 찾기
- 데이터 노이즈에 강건한 표면 이상 탐지 모델 개발  
: SimCLR[Google Research, 2020] 을 이용해 학습한 뒤, feature vector들의 KNN distance 평균을 구해 평균 이상이면 결함으로 탐지 =>

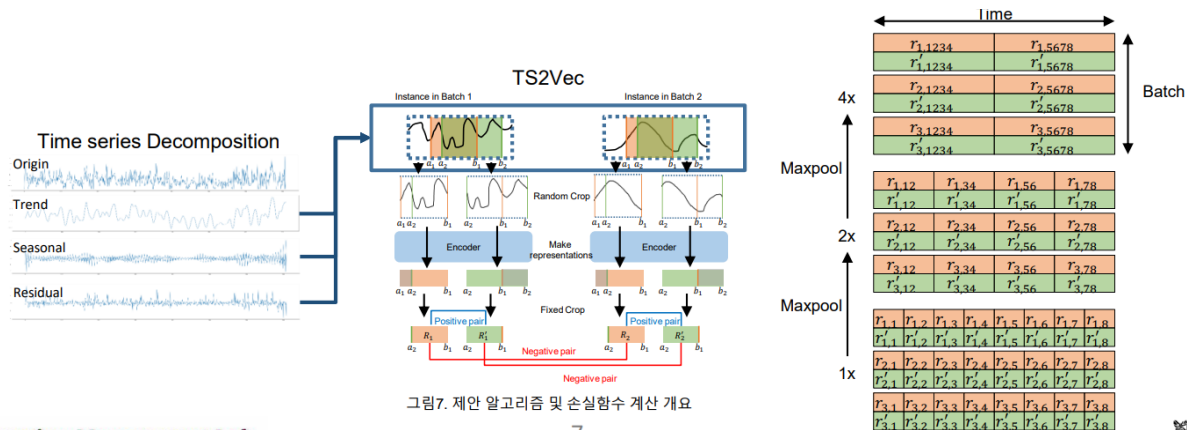
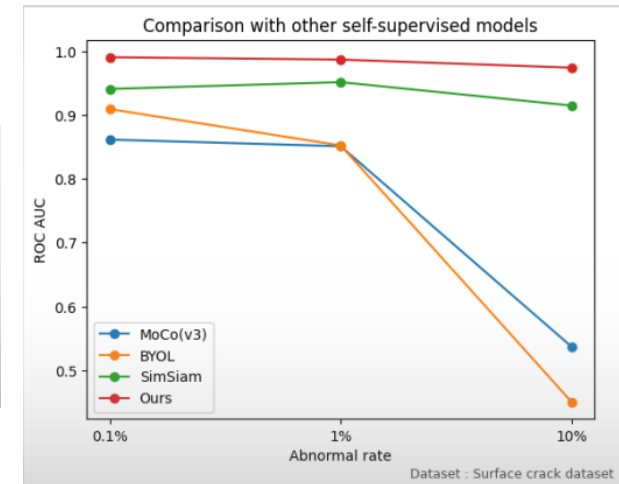


그림7. 제안 알고리즘 및 손실함수 계산 개요



## 이상탐지서비스 알고리즘 소개 및 기업 적용 사례

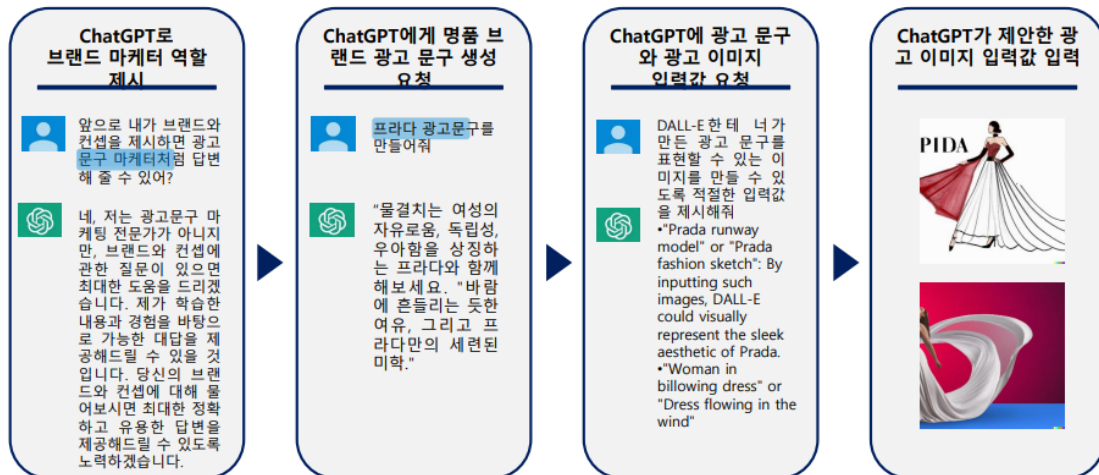
### 이상탐지서비스 알고리즘 소개 및 기업 적용 사례





이재훈 연구위원  
LG전자 AI빅데이터담당  
jae.h.lee@lge.com

- <https://drive.google.com/file/d/1F7ysmBHcbbSKPpBJeWfl-A7aZCLk7ObQ/view?usp=sharing>

## ChatGPT x DALL-E 광고 창작물과 사람의 광고창작물이 고객의 구매행동에 미치는 영향 비교 : 명품 브랜드를 중심으로

- 인공지능 기술의 활용 가능성을 파악하는 데에 의의를 가지며 이를 통해 인공지능 기술이 인간의 창의성을 대체할 수 있는지에 대한 논의 제안
- 사람의 창작물이나 인공지능의 창작물 중 하나만 사용하는 것보다는 두 가지를 협업하거나 균형점을 찾아 결합한 결과물이 더 효과적일 것으로 판단



	Louis Vuitton	Chanel
Original	 <p>L V the truth 루이비통이 진실이다.</p>	 <p>To be irreplaceable, one must be different. 대체할 수 없는 존재가 되기 위해서는, 다른 것이어야 한다.</p>
DALL-E	 <p>Elevate your style with Louis Vuitton's timeless classics. 루이비통의 영원한 클래식 제품으로 스타일을 높이세요</p>	 <p>Experience the timeless elegance of Chanel. 샤넬의 영원한 우아함을 경험하세요.</p>



## 반도체 제조 도메인 특화 지식 그래프 구축을 위한 엔티티 추출 방법 연구

- 반도체 제조 도메인 지식 그래프 구축을 위해 도메인 특화 엔티티(Domain-specific entities)를 자동으로 추출할 수 있는 방법을 탐색 (BERT, SciBERT, ChatGPT, ChatPDF)
  - 2020년~2022년에 발행된 반도체 제조 도메인 특허의 초록에서 엔티티를 추출하고 그 결과를 비교 • 분석
    - BERT 기반 NER 모델은 대문자로 표기된 단어를 위주로 추출함 [축약어 (e.g., UV-LIGA, FinFET), 수학 기호 (e.g.,  $\alpha$ ) 등]
      - FinFET과 같이 대문자로 표기된 엔티티는 잘 추출하나 “wafer”와 같이 소문자로 표기된 보통 명사 엔티티는 잘 추출하지 못함
    - SciBERT 기반 NER 모델은 소문자로 표기된 보통 명사를 위주로 추출
      - “wafer alignment”와 같은 반도체 제조 도메인 특화 엔티티를 잘 추출하지 못함
    - ChatGPT는 NER 모델이 추출하지 못했던 반도체 제조 도메인 특화 엔티티(e.g., wafer alignment)를 잘 추출했다는 점에서 ChatGPT가 NER 모델보다 나은 성능을 보임
      - 추상적 범위의 엔티티(e.g., wafer)를 추출하기도, 구체적 범위의 엔티티(e.g., wafer alignment, photoresist exposure compensation)를 추출하기도 함
    - ChatPDF는 길이가 긴 텍스트를 Input으로 받아들일 수 있다는 장점이 있으나 이를 효과적으로 연산하지는 못함
      - ChatPDF가 추출하는 키워드는 전체 문단이 아닌 소수 개의 문단에서만 등장하는 경향이 있음
      - 전체 문서에서 쿼리와 가장 관련성이 높은 몇 개의 문단을 찾은 후 이를 기반으로 답변을 생성하는 것으로 확인됨
- ➔ 보다 정확한 검증을 위해 본 연구를 통해 추출한 엔티티 목록을 도메인 전문가와 함께 검토해볼 필요가 있음
- ➔ 얼마나 구체적인 범위로 엔티티를 추출할지에 대한 논의가 필요함



## Large Language Model for Healthcare : Severity classification in Triage

- 응급실 환자 데이터를 LLM(GPT) 모델에 넣어서 응급환자인지 판별할 수 있는 모델 개발
- 인풋은 테이블 컬럼값과 그에 해당하는 값을 차례로 입력
- fine-tuning 및 테스트 결과, 다음과 같은 성능을 보임

- Example Prompts

- 'input: Sex 0 Age 71 Arrival mode 0 Injury 0  
Chief\_complain 0 Mental 1 Pain 1 NRS\_pain 2 SBP 160  
DBP 100 HR 84 RR 18 BT 36.6 Saturation 100 Length of  
stay\_min 86 KTAS duration\_min 5.0\noutput: 4',
- 'input: Sex 1 Age 56 Arrival mode 0 Injury 0  
Chief\_complain 1 Mental 1 Pain 1 NRS\_pain 2 SBP 137  
DBP 75 HR 60 RR 20 BT 36.5 Saturation 0 Length of  
stay\_min 64 KTAS duration\_min 3.95\noutput: 2'

### GPT-3 Model

Classifier	Accuracy
Expert	0.885
Nurse	0.895

### Decision Tree Classifier

Classifier	Accuracy
Expert	0.399
Nurse	0.315