

# 제조분야 영상 기반 이상치 탐지 기술 동향

과학기술연합대학원대학교 | 이원준  
한국과학기술원 | 이재영

## 1. 서 론

이상치 탐지 기술은 제조 분야를 포함해 보안 CCTV, 의료 영상, 금융업 등 다양한 분야에서 이상치, 특이치, 또는 비정상 샘플을 구별하는 문제를 말한다. 특히, 제조분야에서는 완성된 제품이 불량인지 양품인지를 구별하는 문제에 많이 사용되고 있으며, 완성된 제품의 영상을 카메라로 취득한 후, 영상에서 이상치가 있는지 없는지를 구별함으로써, 제품의 양불 판정을 내린다.

최근, 딥러닝의 발전과 함께 다양한 딥러닝 기술이 영상 기반 이상치 탐지 기술에 활발히 적용되고 있다. 딥러닝 기반 이상치 탐지 방법은 학습에 이상치 데이터의 사용 여부 및 데이터의 라벨링 여부에 따라 크게 크게 지도학습, 준지도 학습, 비지도 학습으로 구분된다. 이 세가지 접근 중, 이상치 탐지 방법은 지도 학습 접근보다는 준지도 또는 비지도 학습 접근으로 연구가 활발히 되고 있다. 지도 학습 접근의 한계점으로는 첫번째, 이상치 데이터는 정상 데이터에 비해 극히 드물게 발생하므로, 수집하기 어렵다는 점이다. 두번째, 데이터에 라벨을 부여하는 작업은 상당한 시간과 비용을 필요로 한다. 세번째, 이상치 유형을 사전

에 예측하기 어려워 지도 학습 접근으로 유연한 대응이 어렵다는 점이다. 이러한 이유로 이상치 탐지 방법은 준지도 또는 비지도 학습 접근 방식으로 연구가 활발히 되고 있으며, 본 논문에서는 준지도 또는 비지도 학습 접근의 영상 기반 이상치 탐지 기술의 동향을 파악해 보고자 한다.

동향을 파악하기 앞서, 이상치의 정의, 넓은 범위에서 활용 분야를 먼저 소개하고, 통계적 접근, 머신러닝을 이용한 접근, 최근 딥러닝을 이용한 접근에 대해 소개하고자 한다.

## 2. 이상치 탐지 정의 및 유형

일반적으로 이상치는 정상 범주에 속하지 않는 데이터로 정의된다. Hawkins에 의하면 이상치는 “다른 메커니즘에 의해 생성된 것으로 의심되는 정도로 다른 관측치와 크게 다른 관측치”[1]라고 정의했으며, Harmeling에 의하면 “실제 확률 밀도가 매우 낮은 데이터 포인트”[2]로 정의된다. 즉, 이상치는 정상 범주와는 다른 값을 의미하며, 정상에 대한 정의는 각 분야, 문제에 따라 다르게 정의될 수 있다. 이상치는 국문으로는 이상, 이상값, 극단값, 예외, 비정상 등으로

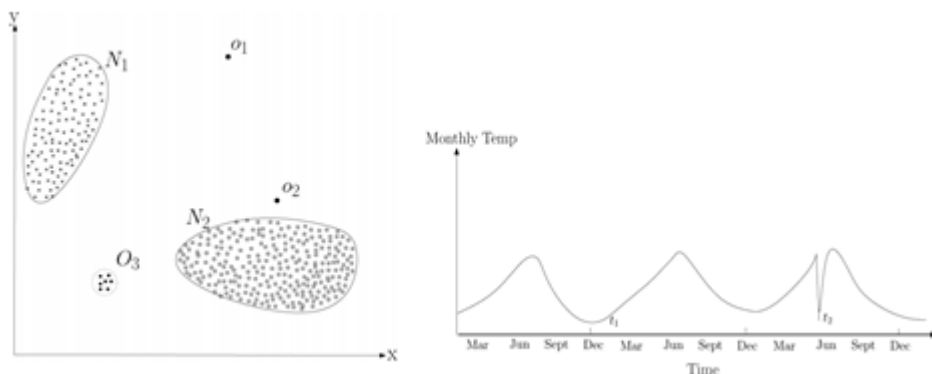


그림 1 이상치의 종류 [7]

부르고, 영문으로는 Anomalies, Outliers, Novelities Exceptions와 같이 표현한다.

이상치 탐지는 이러한 이상치들을 탐지하는 것을 의미하며, 다양한 분야에서 중요한 역할을 수행한다. 이상치 탐지의 주요 목적은 데이터에서 특이한 패턴이나 이벤트를 식별하여 문제를 조기에 발견하고 적절히 대응할 수 있도록 하는 것이다. 일반적으로 이상치는 다음과 같이 세 가지 유형으로 분류된다.

점 이상(Point Anomalies): 단일 데이터 포인트가 예상 패턴, 범위 또는 정상 범주에서 벗어난 경우이다. 점 이상은 일반적으로 무작위로 발생하며 특별한 해석이 없을 수 있다.

집단적 이상(Collective, Group Anomalies): 데이터 내에서 복수의 개체가 전체 데이터 패턴과 다른 경우이다. 집단적 이상은 데이터 포인트 그룹이 전체 데이터 집합과 다른 동작 패턴을 보이는 상황을 나타낸다. 이러한 종류의 이상치는 특정 데이터 포인트 그룹 간의 연관성을 분석하여 찾아낼 수 있다.

맥락적 이상(Contextual, Conditional Anomalies): 특정 상황에서 데이터 집합의 일부가 다른 경우이다. 연속적인 변화의 패턴, 상황적 및 행동적 특징을 모두 고려하여 식별된다. 일반적으로 사용되는 상황적 특징은 시간과 공간이다. 반면 행동적 특징은 소비 패턴, 시스템 로그 이벤트의 발생 또는 정상적인 행동을 설명하는 데 사용되는 모든 특징일 수 있다.

제조 분야 영상 기반 이상치 탐지에서는 완성된 제품의 양불을 판정하는 문제로 대부분의 문제는 맥락적 이상치를 찾는 것으로 볼 수 있다.

### 3. 이상치 탐지 활용 사례

이상치 탐지는 다양한 분야에서 중요한 역할을 수행하고 있으며 의료, 금융, 보안, 제조 산업 등에서 활발하게 연구 및 개발이 진행되고 있다. 이상치 탐지의 중요성과 다양한 활용 분야를 이해하기 위해 다음과 같은 구체적인 예시를 들어 살펴보고자 한다.

의료 분야에서는 이상치 탐지를 통해 질병의 조기 발견이 가능하다. 병원 데이터에는 환자의 나이, 성별, 증상, 검사 결과 등 다양한 유형의 데이터에 다양한 정보가 포함되어 있으며, 이러한 데이터를 이용하여 질병의 위험도나 발병 가능성을 예측할 수 있다. Kauchec[3]등은 심장병을 예측하기 위해 심전도 데이터에 이상치 탐지 알고리즘을 적용하여 심장 박동의 변화를 감지하는 연구를 수행했다.

금융 분야에서는 이상치 탐지를 이용해 이상 거래

를 탐지하여 사기를 예방할 수 있다. 신용 카드 거래 데이터에는 거래 시간, 금액, 장소, 카드 소유자 등 다양한 정보가 포함되어 있으며, 이러한 데이터를 이용하여 이상치 탐지 알고리즘을 적용하면 부정 사용이나 도난 등의 비정상적인 거래를 식별할 수 있다. Malhotra [4] 등은 LSTM(Long Short-Term Memory) 알고리즘을 이용하여 시계열 데이터를 분석하고 이상 거래를 찾아냈다. 이를 통해 신용 카드 부정사용이나 도난 등의 비정상적인 거래를 식별하였다.

보안 분야에서는 네트워크 침입, 악성 코드, DDoS 공격 등 다양한 유형의 위협을 감지하고 대응하기 위해 이상치 탐지가 활용된다. Tang [5]등은 소프트웨어 정의 네트워킹(SDN) 환경에서 Auto-Encoder를 사용하여 네트워크 트래픽 데이터의 특징을 학습하여 이상치를 탐지하였다. 또한, Kim [6] 등은 시스템 로그 데이터를 분석하여 Random Forest 알고리즘을 기반으로 악성코드에 의한 이상치를 탐지하는 연구를 수행하였다. 더불어, Javaid [7]등의 연구에서는 k-평균 클러스터링 알고리즘을 사용하여 네트워크 트래픽 데이터를 군집화하고, 이를 통해 이상 트래픽을 식별하여 네트워크 침입을 탐지하는 방법을 제안하였다. 이러한 방법들은 실시간 보안 모니터링 및 위협 감지에서 효과적으로 활용될 수 있다.

센서 데이터를 기반으로 한 이상치 탐지 알고리즘들 역시 활발히 개발되고 있다. 특히, Janssens [8] 등의 연구에서는 센서 데이터를 활용하여 철강 제조 공정에서의 결함을 탐지하는 방법을 제안하였다. 이 연구는 센서에서 수집된 데이터를 분석하여 제조 공정의 이상 상황을 신속하게 파악하고, 이를 통해 공정의 효율성을 향상시키는 데 초점을 두었다.

제조 산업에서는 이상치 탐지를 결함 검출 및 예측 유지보수에 활용한다. 제조 과정에서 발생하는 센서 데이터나 기계 상태 데이터, 이미지 데이터 등을 이용하여 이상치 탐지 알고리즘을 적용하면 제품의 결함이나 고장을 미리 감지하고 예방하는 데 도움이 된다. Li [9]등은 딥러닝 기반의 지능형 베어링 결함 진단에 관한 연구를 수행하였고, 이를 통해 기계의 고장을 미리 예측하고 관련 부품을 교체할 최적의 시점을 찾는 데 사용했다.

이미지 데이터를 활용한 이상치 탐지 연구도 활발히 진행되고 있다. Dong [10]은 컴퓨터 비전 기술을 사용하여 제조 공정에서 발생하는 이미지 데이터를 분석하였고, 이 연구에서는 딥러닝 기반의 객체 인식 알고리즘을 사용하여 이미지 데이터에서 제품의 결함을 감지한다.

이렇듯 이상치 탐지는 분야나 데이터의 종류에 상관없이 다양한 영역에서 유용하게 활용되며, 이제는 다양한 산업 기술의 필수적인 구성 요소로 자리잡았다. 각각의 분야에서 특정한 문제나 도전 과제를 해결하는 데 이상치 탐지가 큰 역할을 하고 있으며, 이를 통해 성능 향상, 사고 예방, 비용 절감 등 다양한 이점을 얻을 수 있다. 이상치 탐지 기술의 발전과 활용은 앞으로도 계속되며, 그 중요성은 더욱 커질 것으로 예상된다.

#### 4. 라벨링 여부에 따른 이상치 탐지 방법 분류

이상 탐지 분석을 위한 데이터에 정상 혹은 비정상의 정답을 부여하는 작업을 라벨링(Labeling)이라고 한다. 이상치 탐지 방법은 학습 시 이상치 데이터의 사용 여부 및 데이터의 라벨링 여부에 따라 크게 지도학습, 준지도 학습, 비지도 학습으로 구분할 수 있다.

##### 4.1 지도학습 이상치 탐지

지도학습(Supervised Learning)이란 이미 정답이 부여된 데이터로 학습을 하는 기법이다. 즉, 이상치와 정상 데이터를 모두 사용하여 이상치 탐지 학습을 수행한다. 또한 데이터의 라벨링이 되어 있어 다른 방법 대비 정확도가 높은 특징이 있으나 현실에서 라벨링이 된 데이터셋을 구축하기에는 다음과 같은 한계점들이 존재한다.

첫째, 이상치 데이터는 정상 데이터에 비해 극히 드물게 발생하여 수집되므로, 이로 인해 데이터 불균형(클래스 불균형) 문제가 발생해 머신러닝, 딥러닝 성능에 부정적인 영향을 미친다. 이 문제를 해결하기 위해 과대 표본화(Over Sampling), 과소 표본화(Under Sampling), SMOTE 등의 기법이 사용되지만, 이들 기법은 모델의 과적합 문제를 야기할 수 있다.

둘째, 데이터에 라벨을 부여하는 작업은 상당한 시간과 비용을 필요로 한다. 이로 인해 고품질의 라벨링된 데이터셋을 구축하는 것은 어렵다. 또한, 라벨링

과정에서 작업자의 주관적 판단이 개입되어 데이터의 품질이 떨어질 수 있다.

마지막으로 이상치의 유형을 사전에 예측할 수 없기 때문에 새로운 유형의 이상치가 발생했을 때 지도 학습의 방식으로는 유연한 대응을 할 수 없다. 이로 인해 지도학습 기반의 이상치 탐지 기법은 한정된 유형의 이상치가 발생하는 환경에는 효과적일 수 있지만, 이외의 환경에서는 효과적이지 않다.

이러한 한계점들로 인해 지도학습 이상치 탐지 방법보다는 비지도, 준지도 학습 방법에 대한 연구가 활발히 진행되고 있다.

##### 4.2 비지도, 준지도 학습 이상치 탐지

비지도 학습(Unsupervised Learning)은 지도학습과 달리 정답이 없는 데이터를 이용하여 데이터의 구조나 패턴을 탐색하는 학습 방법이다. 이상치 탐지에서 비지도 학습은 라벨링 되지 않은 데이터를 사용해 데이터의 특징과 패턴을 이해하고, 이를 바탕으로 이상치를 탐지한다.

준지도 학습(Semi-supervised Learning)은 라벨링된 데이터와 라벨링되지 않은 데이터를 모두 사용하는 학습 방법이다. 이상치 탐지에서의 준지도 학습은 라벨링된 정상 데이터를 주로 사용하며, 라벨링된 이상치 데이터는 사용하지 않는다. 이 방법은 지도학습의 정확도와 비지도 학습의 유연성을 결합하여, 라벨이 부족한 데이터에서도 이상치 탐지를 수행할 수 있다.

비지도 및 준지도 학습 기반의 이상치 탐지 방법은 라벨이 없거나 부족한 데이터에서도 작동할 수 있다. 이러한 방법들은 정상 데이터의 패턴을 학습하고, 이를 기반으로 이상치를 탐지한다. 따라서, 지도학습이 가지는 한계점들을 극복할 수 있다.

또한, 이러한 방법들은 새로운 유형의 이상치에 대한 유연한 대응이 가능하다. 지도학습 방식은 새로운 이상치 유형에 대응하기 어렵지만, 준지도 및 비지도 학습 방식은 데이터의 패턴과 구조를 학습하므로, 새로운 유형의 이상치가 발생하더라도 그 패턴을 학습하여 탐지할 수 있다.

표 1 학습 데이터에 따른 이상치 방법 분류

학습 방법	정상 데이터	이상치 데이터
지도학습 (Supervised Learning)	학습에 사용	학습에 사용
준지도 학습 (Semi-supervised Learning)	학습에 사용	학습에 사용하지 않음
비지도 학습 (Unsupervised Learning)	학습에 사용하지 않음	학습에 사용하지 않음

## 5. 이상치 탐지 기법

### 5.1 통계적 이상치 탐지 기법

통계적 이상치 탐지 방법은 데이터의 통계적 특성을 분석하여 이상치를 식별하는 방법이다. 이 방법은 데이터가 정규 분포를 따른다는 가정하에 작동하며, 평균과 표준편차와 같은 통계적 측정치를 사용하여 이상치를 식별한다.

통계적 이상치 탐지 방법 중 하나는 Z-점수 방법이다. Z-점수는 각 데이터 포인트가 평균에서 얼마나 떨어져 있는지를 측정하는 통계적 측정치이다. Z-점수가 일정 임계값보다 큰 데이터 포인트는 이상치로 간주될 수 있다.

다른 방법으로 IQR(Interquartile Range) [11]이 있다. IQR은 제 1사분위수와 제 3사분위수 사이의 범위로, 데이터의 중앙값을 기준으로 분포의 전반적인 형태를 나타낸다. IQR의 1.5배 이상 떨어진 데이터 포인트는 이상치로 간주될 수 있다.

통계적 이상치 탐지 방법은 간단하고 계산이 빠르지만, 데이터가 정규 분포를 따르지 않거나 다차원 데이터에는 적용하기 어려울 수 있다. 따라서 특정 상황과 데이터 유형에 따라 적합한 방법을 선택하는 것이 중요하다.

### 5.2 머신러닝 이상치 탐지 기법

머신러닝 이상치 탐지 방법에는 다양한 접근법이 있다. 그 중 하나는 밀도 기반 이상치 탐지이다. 이들은 크게 두 가지 알고리즘으로 분류할 수 있다. 첫 번째는 k-최근접 이웃[12,13]으로 정상 데이터 포인트가 밀집된 이웃 주변에 나타나고 이상치는 멀리 떨어져 있다고 가정한다. 가장 가까운 데이터 포인트 집합은 유클리디안 거리 혹은 맨해튼, 민코프스키, 햄밍 거리와 같은 거리 지표에 기초하여 데이터를 분류하는 간단하고 비모수적인 학습 기법이다.

두 번째는 데이터의 상대 밀도로 측정하는 방법으로, 지역 이상치 계수(LOF) [14]로 더 잘 알려져 있다. 이 방법은 k-최근접 이웃 알고리즘을 통해 밀도를 추정하고, 도달 가능성 거리를 기반으로 이웃 밀도와 비교한다. 이상치는 상대적으로 낮은 밀도를 가지며, LOF 값이 1보다 큰 경우로 판단된다.

다른 방법으로 군집 기반 이상치 탐지가 있다. 이 방법은 유사한 데이터 포인트가 지역 중심점으로부터의 거리에 따라 유사한 그룹 또는 군집에 속한다는 것으로 가정한다. 가장 대표적인 알고리즘인 k-means [15]는 광범위하게 사용되는 군집화 알고리즘이며, 데

이터 포인트의 'k'개의 유사한 군집을 생성한다. 이러한 군집의 특징과 다른 데이터 인스턴스는 이상치로 측정한다.

마지막으로, 서포트 벡터 머신(SVM) [16] 기반 이상치 탐지가 있다. SVM은 일반적으로 지도 학습과 관련이 있지만, One Class SVM과 같은 확장을 사용하여 라벨이 없는 훈련 데이터를 포함하는 비지도 문제로서 이상을 식별할 수 있다. 알고리즘은 정상 데이터를 군집화하기 위해 부드러운 경계를 학습한 다음, 테스트 데이터를 사용하여 학습된 영역 외부에 있는 이상치를 식별하도록 조정한다.

### 5.3 딥러닝 기반 이상치 탐지 기법

딥러닝(Deep Learning)은 여러 개의 은닉 계층을 갖는 신경망을 사용하여 많은 양의 데이터와 복잡한 패턴을 처리하는 데 뛰어난 성능을 발휘하며 최근 몇 년 동안 다양한 분야에 빠르게 적용되어 왔다. 이러한 발전에 따라 이상치 탐지에도 딥러닝 기반 기법들이 적용되어, 복잡한 패턴과 구조를 갖는 데이터에서 효과적으로 이상치를 탐지할 수 있는 연구가 활발히 진행되고 있다.

이 장에서는 One-Class Classification에 초점을 맞춰 딥러닝 기반의 주요 이상치 탐지 방법들을 소개한다. 특히, 재구성 기반 방법(Reconstruction-based Method)과 임베딩 유사도 방법(Embedding-similarity based Method)에 대해 자세히 알아본다. 마지막으로, 현장 상황을 고려한 최신 연구에 대해 소개한다.

#### 5.3.1 재구성 기반 방법(Reconstruction-based Method)

재구성 기반 이상치 탐지 방법은 정상 데이터만 학습한 모델이 정상 데이터를 잘 재구성할 것이라는 가정하에, 재구성 오류를 기반으로 이상치를 탐지하는 방법이다. 대표적으로 오토 인코더(Auto Encoder, AE), 변이형 오토 인코더(Variational AutoEncoder, VAE), 적대적 생성 모델(Generative Adversarial Network, GAN) 등이 있다. 이들 방법은 픽셀 수준에서의 이상치 탐지 능력이 뛰어나며, 사전 학습된 모델 없이 처음부터 학습한다.

오토 인코더(AE) [17] 기반 방법은 입력 데이터를 압축하여 저차원 표현으로 만든 후, 이를 복원하는 과정에서 데이터의 재구성 오류를 최소화하도록 학습한다. 정상 데이터의 패턴을 학습하며, 이상치는 재구성 오류가 크게 나타날 것으로 예상하여 이상치를 탐지한다. 오토 인코더는 입력 데이터를 압축하는 인코더와 압축된 데이터를 복원하는 디코더로 구성된다. 인코더는 입력을 잠재 변수로 압축하고, 디코더는 잠재

변수를 원 데이터로 복원한다. 이를 통해 데이터 레이블이 없이도 정상 상태의 주요 특징을 학습한다[18]. 학습이 완료된 오토 인코더에 정상 상태 데이터를 입력하면, 디코더에서 입력과 동일한 데이터가 출력된다. 반면, 비정상 상태 데이터가 입력되면, 인코더와 디코더가 정상 데이터만으로 학습했기에 정확한 압축과 복원을 수행하지 못한다. 따라서, 디코더가 출력하는 최종 데이터는 최초 입력 데이터와 다르게 복원된다. 결과적으로, 오토 인코더를 사용하여 입력 데이터와 출력 데이터의 차이, 즉, 재구성 손실을 계산하여 이를 통해 입력 데이터가 정상 데이터인지 비정상 데이터인지 판단하게 된다. 이 기법은 재구성 손실을 기반으로 이상치를 탐지하는데 활용된다.

변이형 오토 인코더(VAE) [19]는 오토 인코더(AE)의 확장된 형태로, 입력 데이터를 압축하고 복원하는 과정에서 확률적 접근 방식을 사용한다. VAE는 인코더, 잠재 변수, 디코더의 세 부분으로 구성된다. 인코더는 입력 데이터를 잠재 변수의 확률 분포로 압축한다. VAE의 중요한 특징 중 하나는 인코더가 잠재 변수의 확률 분포를 추정하는 과정에서 평균과 분산을 고려한다. 이를 통해 인코더는 다양한 확률 분포를 학습할 수 있다. 디코더는 잠재 변수를 원 데이터로 복원한다. VAE의 핵심은 인코더에서 출력되는 잠재 변수의 확률 분포를 사용하여 학습 및 데이터 복원 과정에서 노이즈를 주입한다는 점이다[20]. 이를 통해 VAE는 데이터의 불확실성을 포착하고, 복원 과정에서 다양한 데이터를 생성할 수 있다.

학습이 완료된 VAE에 정상 상태 데이터를 입력하

면, 디코더에서 입력과 유사한 데이터가 출력된다. 그러나 비정상 상태 데이터가 입력되면, 인코더와 디코더가 정상 데이터의 확률 분포를 기반으로 학습되었기에 정확한 압축과 복원을 수행할 수 없다. 그러므로, 디코더가 출력하는 최종 데이터는 입력 데이터와 상이하게 된다.

AE와 VAE는 모두 재구성 손실을 이용하여 이상치를 탐지하지만, VAE의 경우 확률적 접근 방식을 사용하여 데이터의 불확실성을 포착하는 것이 주요 차이점이다. 결과적으로, VAE를 사용하여 입력 데이터와 출력 데이터의 차이, 즉 재구성 손실을 계산하여 입력 데이터가 정상 데이터인지 비정상 데이터인지 판단할 수 있다. 이 기법은 재구성 손실을 기반으로 이상치를 탐지하는데 활용된다[21].

생성적 적대 신경망(GAN) [22]은 Ian Goodfellow가 2014년 발표한 이후 다양한 분야에서 기반 연구가 되어 왔다. GAN은 생성자와 판별자로 구성된 신경망으로, 생성자는 실제 데이터와 유사한 가짜 데이터를 생성하고, 판별자는 실제 데이터와 가짜 데이터를 구별하도록 학습한다. 생성자는 적대적 학습을 통해 점차 실제와 구별하기 어려운 가짜 데이터를 생성해내며, 판별자는 실제 샘플과 가짜 샘플을 구별하는 능력을 향상시킨다. 이러한 과정을 통해 생성자는 점차 더 정교한 데이터를 생성하게 된다. 생성자 네트워크는 고정 차원의 임의의 노이즈로 벡터부터 입력 데이터 분포의 구성 요소와 유사한 샘플을 만들어내는 방법을 학습한다. 판별자는 생성자가 생성한 가짜 샘플과 원본 데이터에서 추출된 실제 샘플을 정확하게 구별하

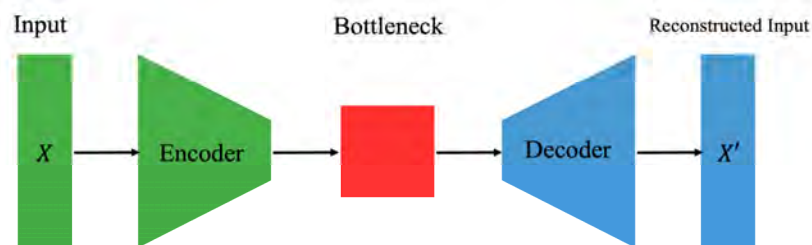


그림 2 Auto-Encoder 구조

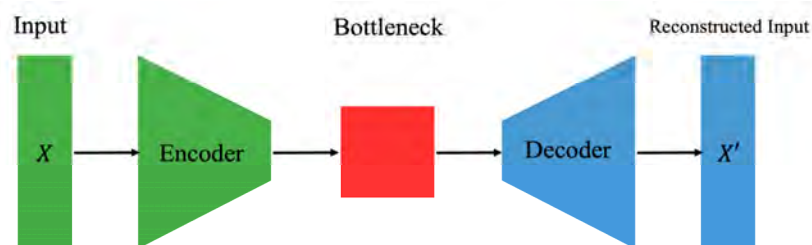


그림 3 Variational Auto-Encoder 구조

는 방법을 학습한다. 훈련 과정에서 각 학습마다 생성자의 매개 변수는 판별자가 구별할 수 없게 업데이트되며, 판별자의 매개 변수는 생성된 샘플과 실제 샘플을 정확하게 구별할 수 있는 능력을 극대화하도록 업데이트된다. 학습이 진행됨에 따라 생성자는 원본 데이터와 유사한 샘플을 효과적으로 생성하는 데 능해지고, 판별자는 실제와 가짜 샘플을 구분하는 작업에 대한 능력을 향상시킨다.

재구성 기반 이상치 탐지에 GAN을 사용할 경우, 생성자는 정상 데이터의 분포를 학습하고, 판별자는 생성자가 만든 데이터와 실제 데이터를 비교하여 이상치를 탐지하게 된다.

### 5.3.2 임베딩 유사도 방법(Embedding-similarity based Method)

이상치 탐지에 있어 임베딩 유사도 기반 방법론은 핵심적 역할을 수행한다. 본 방법론은 이미지 임베딩 또는 이미지 패치 임베딩을 생성하여 임베딩 공간에서 학습 데이터와 테스트 데이터 간의 유사도를 기반으로 이상치를 식별한다. 이미지 임베딩은 심층 신경망을 통해 이미지를 벡터 형태로 표현한 것이며, 이미지 패치 임베딩은 이미지의 일부인 패치 단위로 이미지 패치 임베딩을 생성하며 주로 Anomaly localization에 사용된다. 이러한 임베딩 유사도 기반 방법론은 딥러닝 모델을 학습하는 데 필요한 대량의 고품질 데이터셋의 요구를 줄일 수 있다는 점에서 지도학습의 한계를 극복한다. 또한 ImageNet [23], CIFAR10 [24] 같은 대규모 학습 데이터셋으로 사전 훈련된 모델의 가중치를 활용하여 이상치를 탐지하므로, 사전 훈련에

대한 요구가 없어 효율적이며 추론 속도 및 정확도가 향상된다.

최근 연구들, 특히 SPADE [25], Mahalanobis AD [26], PaDiM [27], PatchCore [28]은 임베딩 유사도 기반 방법론을 사용하여 유의미한 결과를 도출해냈다. 이들 논문은 산업 제조 결함 데이터셋인 MvTec AD [29]를 벤치마크로 사용하여 연구가 진행되었다.

SPADE는 이미지의 하위 영역(sub-image)에서 픽셀 단위로 이상 탐지를 수행하는 알고리즘이다. 이 알고리즘은 딥러닝 기반의 피라미드 특징 상관성(Deep pyramid correspondence)을 사용하여 피라미드 방식으로 특징을 추출하고, 특징 간의 상관관계를 비교함으로써 이상치를 탐지한다.

ImageNet 데이터로 사전 학습된 Wide-ResNet50을 사용하고, 마지막 합성곱 계층을 전역 풀링한 후 얻은 특징 벡터를 사용하여 특징을 추출한다. 테스트 이미지가 주어지면, K-최근접 이웃(KNN) 알고리즘을 사용하여 학습 데이터 세트에서 가장 유사한 K개의 정상 이미지를 검색한다. 이 과정에서 유클리디안 거리를 사용하여 이미지 간의 거리를 측정하며, 이를 통해 이미지 수준에서 이상치를 판별한다. 이미지 수준에서 이상으로 판별된 경우, 다음 단계는 이미지의 픽셀 수준에서 이상을 탐지하는 것이다. 이를 위해, 먼저 테스트 이미지와 가장 유사한 K개의 정상 이미지를 피라미드 특징 매칭 방식(Feature Pyramid Matching)을 통해 정렬하고, 이를 통해 이상 점수를 계산한다. 픽셀의 이상 점수가 특정 임계값을 초과하는 경우, 해당 픽셀을 이상으로 판단한다. 이때, ResNet에서 추출

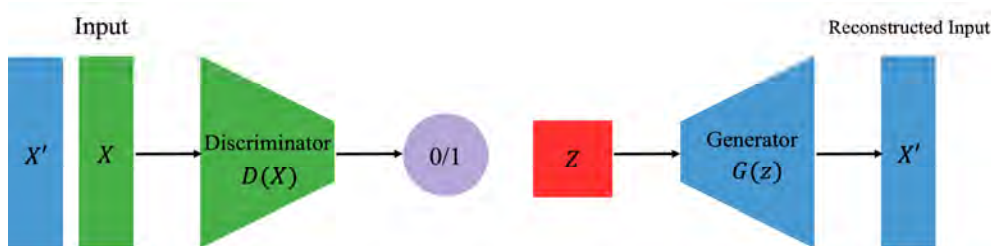


그림 4 Generative Adversarial Network 구조

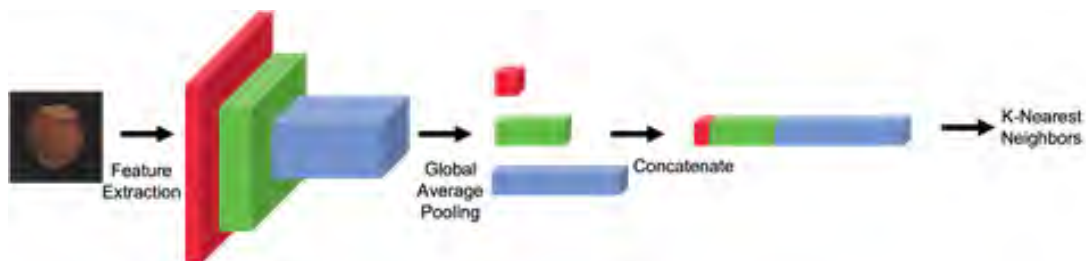


그림 5 SPADE 학습 과정



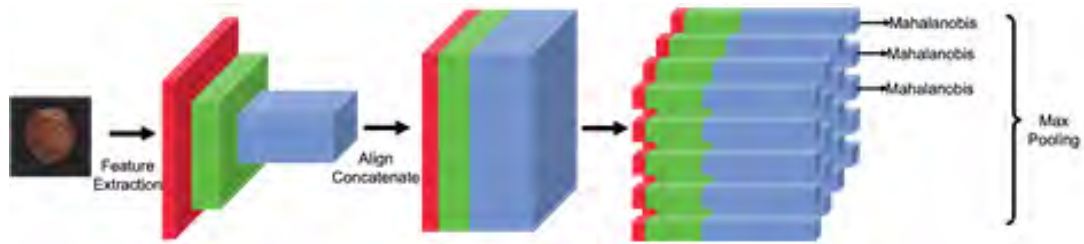


그림 6 PaDiM 학습 과정

한 여러 레벨의 피라미드 특징을 이용한다. 이 특징들은 세부적인 로컬 특징과 글로벌 컨텍스트를 모두 반영하며, 이미지의 특징을 더욱 정확하게 파악하는 데 도움이 된다.

SPADE의 핵심 아이디어는 딥러닝 기반 특징 추출과 전통적인 최근접 이웃 방법을 결합하여 이상치 탐지를 수행하는 것이다. 그러나 SPADE의 KNN을 사용한 접근 방식은 계산 복잡도가 선형적으로 증가하는 한계를 갖는다. 이로 인해 산업 이상 탐지에서의 효율성이 떨어진다. 이러한 한계를 극복하기 위해 Mahalanobis AD가 연구되었다.

Mahalanobis AD는 사전 훈련된 모델에서 추출한 특징들의 분포를 모델링하여 이상치를 탐지하는 방법을 제안하였다. 이 방법은 각 이미지 특징이 다변량 가우시안 분포를 따른다고 가정하고, 이상치와 정상 데이터 간의 거리를 mahalanobis 거리를 사용하여 측정한다. 이렇게 하여 KNN의 계산 복잡도 문제를 해결하였다. 하지만 이 방법은 패치 수준이 아닌 이미지 수준에서만 작동하여 성능 향상을 극대화하지 못한다.

SPADE와 Mahalanobis AD 두 방법론의 한계를 극복하기 위해, PaDiM은 다변량 가우시안 모델링을 활용한 새로운 이상치 탐지 방법을 제안한다. PaDiM은 SPADE의 이미지 패치 단위의 이상치 탐지 기능과 Mahalanobis AD의 mahalanobis 거리를 활용한 이상치 측정 방식을 결합하였다. 이렇게 함으로써, 복잡한 KNN 방식 대신 Mahalanobis 거리를 이용하여 이상치를 더욱 효율적으로 측정함과 동시에, 패치 수준으로

이상치를 탐지해 이미지 수준 대비 좀 더 정밀한 결과를 보인다.

PatchCore [29]는 PaDiM 방법론 기반으로 개발된 방법론이다. PaDiM은 encoder로 부터 얻은 특징점들을 random하게 선택해서 추론에 활용하는 반면, PatchCore는 encoder로부터 얻은 특징점들이 고차원 공간상에서 최대한 넓게 퍼져있도록 특징점들을 샘플링해 memory bank에 저장해두고, 추론에 활용한다. 특징 공간의 공간적 특징을 활용함으로써, PatchCore는 PaDiM 대비 더 나은 성능을 보인다.

### 5.3.3 현장 상황을 고려한 딥러닝 기반 이상치 탐지 연구

위에 소개한 방법론은 one-class classification 방법론으로, 하나의 class마다 학습을 진행하고 추론에 활용해야 한다. 하지만, class 별로 학습 weight를 저장하는 것은 메모리 관점에서 비효율적이고, 학습된 weight를 관리하는 것 또한 어려운 문제이다. 이러한 문제를 극복하고자, 기존 one-class classification으로 이상치 탐지 문제를 해결한 것과 달리, UniAD(Unified Anomaly Detection)[30]은 multi-class anomaly detection 방법론을 제시한다. UniAD는 임베딩 유사도 기반 방법과 같이 encoder를 통해 특징점을 추출하되, 이 특징점 일부를 마스킹해 원 상태로 복원시키는 재구성 모듈로 구성되어 있다. UniAD는 one-class classification 방법론의 성능에 거의 근접한 성능을 보이며, 현장 상황에서 메모리 및 학습 모델 관리 측면에서 이점을 제공하는 방법론을 제시한다.

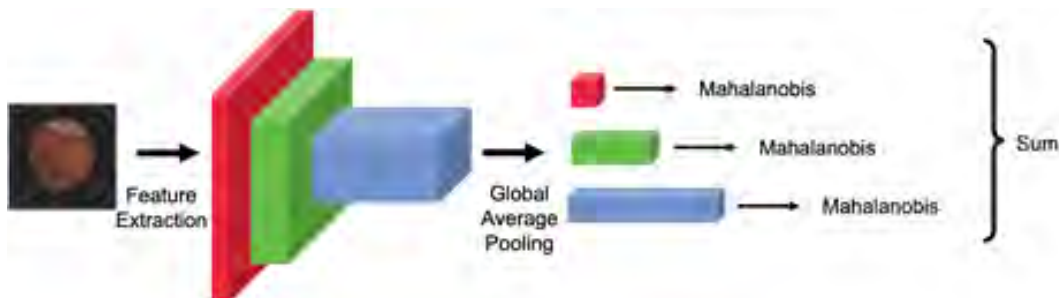


그림 7 Mahalanobis AD 학습 과정

또 다른 현장 상황을 고려한 연구로는 few-shot 이상치 탐지 연구가 있다. 새로운 제품을 개발하는 경우 특히, 양산 과정에 들어가기 전에는, 정상 샘플의 영상도 확보하기 어렵다. 이와 같이 정상 샘플 수도 확보하기 어려운 상황에 적용할 수 있는 다양한 few-shot 이상치 탐지 기술이 최근 많이 제안되고 있다. 특히, RegAD(Registration based Anomaly Detection)[31]은 사람이 참고할 만한 몇장의 예시를 살핌으로써 그 예시를 기준으로 불량을 판별해내는 과정을 모방하여 알고리즘을 제안한다. RegAD는 정합 모듈, encoder, predictor 세개의 모듈로 구성되어 있으며, 이 구성이 병렬적으로 Siamese network 구조로 이루어져 있다. RegAD에서 한 branch는 몇 장의 정상 샘플을 입력으로 받고, 다른 branch는 test 영상을 입력받는 것과 같이 구성되어, 다양한 class에 대해 이러한 상황을 network가 학습하도록 한다. 정상 샘플이 제한적인 상황을 가정했기 때문에, 위에서 소개한 one-class classification 접근 방법들에 비해 성능은 다소 떨어지지만, 새로운 모델을 개발할 때, 개발 초기 단계의 현실적인 상황을 고려한 방법으로써 실용성이 높은 방법이다.

## 6. 결 론

본 논문에서는 제조분야 이상치 탐지 기술에 대해 살펴보았다. 특히, 최근떠오르는 딥러닝 기반 방법론 중 준지도 및 비지도 학습 방법론에 해당하는 방법을 중점적으로 살펴보았다. 대부분의 경우, 일반적으로 one-class classification 방법론으로 접근하며, 이 방법론은 크게 재구성 기반 방법과 임베딩 유사도 기반 방법으로 나눌 수 있다. 이러한 방법들과 별개로, 제조 현장 상황을 고려한 multi-class classification과 few-shot anomaly detection 방법으로 최근 제안된 방법도 살펴보았다. 지금까지 제조 분야에서 이상치 탐지 연구는 MVTec AD 데이터셋을 많이 사용해 왔고, MVTec AD 데이터셋에서 one-class classification 접근의 성능은 높은 수준으로 확보가 되었다. 따라서, 최근들어서 one-class classification을 넘어 좀 더 실용적인 문제를 해결하는 방법들이 제안이 되고 있고, 앞으로도 현장 상황을 고려한 실용적인 방법들이 많이 제안될 것으로 보인다.

## 참고문헌

[ 1 ] Hawkins, D. M. (1980). Identification of outliers (Vol. 11). London: Chapman and Hall.

[ 2 ] Harmeling, S., Dornhege, G., Tax, D., Meinecke, F., and Müller, K. R. (2006). From outliers to prototypes: ordering data. *Neurocomputing*, 69 (13-15), 1608-1618.

[ 3 ] Kachuee, M., Kiani, M. M., Mohammadzade, H., & Shabany, M. (2018). ECG heartbeat classification: A deep transferable representation. *IEEE Access*, 6, 24693-24703.

[ 4 ] Malhotra, P., Vig, L., Shroff, G., & Agarwal, P. (2015). Long short term memory networks for anomaly detection in time series. In *Proceedings. Presses universitaires de Louvain*, 89, 89-94.

[ 5 ] Tang et al. (2016). Deep learning approach for network intrusion detection in software defined networking.

[ 6 ] Kim, J., Kim, J., Thu, H. L. T., & Kim, H. (2017). Long short term memory recurrent neural network classifier for intrusion detection. In *2017 International Conference on Platform Technology and Service (PlatCon)*, 1-5.

[ 7 ] Javaid, A., Niyaz, Q., Sun, W., & Alam, M. (2016). A deep learning approach for network intrusion detection system. In *2016 9th International Conference on Communication Software and Networks (ICCSN)*, 110-114.

[ 8 ] Janssens, O., Slavkovikj, V., Vanden Berghe, S., Janssens, K., Vanrumste, B., & Weyn, M. (2019). Convolutional neural network based fault detection for rotating machinery. *Journal of Sound and Vibration*, 377, 331-345.

[ 9 ] Li, B., Zhang, Y., Guo, T., & Li, P. (2018). An intelligent fault diagnosis method using unsupervised feature learning towards mechanical big data. *IEEE Transactions on Industrial Electronics*, 65(5), 3906-3917.

[10] Dong, Y., & Peng, Y. (2019). Anomaly detection in manufacturing systems using structured neural networks. *IEEE Transactions on Industrial Informatics*, 15(6), 3702-3709.

[11] Iglewicz, B., & Hoaglin, D. (1993). Volume 16: how to detect and handle outliers. *The ASQC basic references in quality control: statistical techniques*, 16, 1-85.

[12] M. M. Breunig, H. Kriegel, R. T. Ng, and J. Sander, "LOF: Identifying density-based local outliers". *Proceedings of ACM SIGMOD 2000*.

[13] K. Bache and M. Lichman, "UCI Machine Learning Repository." 2013. <http://archive.ics.uci.edu/ml>.

[14] S. D. Bay and M. Schwabacher. "Mining Distance-Based Outliers in Near Linear Time with Randomization and a Simple Pruning Rule." *Proceedings of the ACM*



SIGKDD 2003.

- [15] J. MacQueen, "Some Methods for Classification and Analysis of MultiVariate Observations", Proceedings of 5th Berkeley Symposium on Mathematical Statistics and Probability, given in 1965. pp. 281-297, 1967.
- [16] B. Schölkopf, J. C. Platt, J. C. Shawe-Taylor, A. J. Smola, and R. C. Williamson "Estimating the Support of a High-Dimensional Distribution", Neural Computation 13(7): 1443-1471, 2001.
- [17] LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. Nature, 521(7553), 436-444.
- [18] Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep Learning. MIT press.
- [19] Kingma, D. P., & Welling, M. (2013). Auto-Encoding Variational Bayes. arXiv preprint arXiv:1312.6114.
- [20] Masci, J., Meier, U., Ciresan, D., & Schmidhuber, J. (2011). Stacked convolutional auto-encoders for hierarchical feature extraction. Artificial Neural Networks and Machine Learning Conference.
- [21] An, J., & Cho, S. (2015). Variational Autoencoder based Anomaly Detection using Reconstruction Probability. SNU Data Mining Center.
- [22] Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., ... & Bengio, Y. (2020). Generative adversarial networks. Communications of the ACM, 63(11), 139-144.
- [23] Deng, J., Dong, W., Socher, R., Li, L. J., Li, K., & Fei-Fei, L. (2009, June). Imagenet: A large-scale hierarchical image database. In 2009 IEEE conference on computer vision and pattern recognition (pp. 248-255). Ieee.
- [24] Krizhevsky, A., & Hinton, G. (2009). Learning multiple layers of features from tiny images.
- [25] Cohen, N., & Hoshen, Y. (2020). Sub-image anomaly detection with deep pyramid correspondences. arXiv preprint arXiv:2005.02357.
- [26] Rippel, O., Mertens, P., & Merhof, D. (2021, January). Modeling the distribution of normal data in pre-trained deep features for anomaly detection. In 2020 25th International Conference on Pattern Recognition (ICPR) (pp. 6726-6733). IEEE.
- [27] Defard, T., Setkov, A., Loesch, A., & Audigier, R. (2021, March). Padim: a patch distribution modeling framework for anomaly detection and localization. In Pattern Recognition. ICPR International Workshops and Challenges: Virtual Event, January 10-15, 2021, Proceedings, Part IV (pp. 475-489). Cham: Springer International Publishing.
- [28] Roth, K., Pemula, L., Zepeda, J., Schölkopf, B., Brox, T., & Gehler, P. (2022). Towards total recall in industrial anomaly detection. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 14318-14328).
- [29] Bergmann, P., Fauser, M., Sattlegger, D., & Steger, C. (2019). MVTec AD--A comprehensive real-world dataset for unsupervised anomaly detection. In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition (pp. 9592-9600).
- [30] You, Z., Cui, L., Shen, Y., Yang, K., Lu, X., Zheng, Y., & Le, X. (2022). A unified model for multi-class anomaly detection. Advances in Neural Information Processing Systems, 35, 4571-4584.
- [31] Huang, C., Guan, H., Jiang, A., Zhang, Y., Spratling, M., & Wang, Y. F. (2022, October). Registration based few-shot anomaly detection. In European Conference on Computer Vision (pp. 303-319). Cham: Springer Nature Switzerland.

## 약 력

### 이 원 준



2021 延边大学科学技术学院 경영정보처리학과 졸업(학사)

2021~현재 한국전자통신연구원 UST학생연구원

2021~현재 과학기술연합대학원대학교 인공지능학과 석사과정

관심분야: 이상치 탐지

Email : wjdjnswns@ust.ac.kr

### 이 재 영



2015 서강대학교 전자공학과 졸업(학사)

2017 서강대학교 전자공학과 졸업(석사)

2017~2020 LG 이노텍

2022 한국전자통신연구원 위촉연구원

2020~현재 한국과학기술원 전기및전자공학과 박사과정

관심분야: 이상치 탐지

Email : mcneato@kaist.ac.kr