

TimeVQVAE-AD를 활용한 시계열 이상 탐지: 성능 검증과 실험 분석

임수환, 최현민, 성민규, *강재모

경북대학교

(tnghks9317, iissaacc, alsrb0351, *jmkang)@knu.ac.kr,

TimeVQVAE-AD for Time Series Anomaly Detection: Performance Validation and Experimental Analysis

Im Su Hwan, Choi Hyun Min, Sung Min Gyu, Kang Jae-Mo*

Kyunpook National Univ.

요약

본 논문은 시계열 데이터 이상 탐지(Time Series Anomaly Detection, TSAD)는 다양한 산업 분야에서 중요한 역할을 하며, 시스템의 안정성과 성능을 유지하는 데 필수적인 기술이다. 최근 제안된 TimeVQVAE-AD는 Masked Generative Modeling을 기반으로 하여 높은 탐지 정확도와 설명 가능성을 동시에 제공하는 혁신적인 방법론으로 주목받고 있다. 본 논문에서는 TimeVQVAE-AD를 소개하고, 해당 방법론의 성능을 독립적으로 검증하기 위해 UCR Time Series Anomaly (UCR-TSA) 아카이브를 활용한 실험을 수행한다.

I. 서론

시계열 이상 탐지(Time Series Anomaly Detection, TSAD)는 다양한 산업 분야에서 중요한 문제로 다뤄지고 있으며, 시스템의 안정성 유지와 조기 이상 탐지를 위해 필수적인 기술로 자리 잡고 있다. 최근 몇 년간 TSAD 분야에서는 다양한 기법들이 제안되었고, 그중 딥러닝 기반 방법 [1]들이 많은 주목을 받았다. 특히, TimeVQVAE-AD[2]라는 새로운 방법론은 Masked Generative Modeling을 기반으로 하여 높은 탐지 정확도와 설명 가능성을 동시에 제공한다.

본 논문에서는 TimeVQVAE-AD를 소개하고, 해당 방법론의 성능을 검증하고자 한다. 이를 위해, 기존 연구에서 사용된 UCR Time Series Anomaly (UCR-TSA)[3] 아카이브를 기반으로 실험을 수행하고, 기존 연구에서 보고된 성능이 실제로 일관되게 재현될 수 있는지 평가한다. 이를 통해 TimeVQVAE-AD의 실제 성능과 설명 가능성을 명확히 이해하고, 이 방법론이 실제 세계 응용에 얼마나 적합한지를 검토한다.

II. 본론

2.1 TimeVQVAE-AD

TimeVQVAE-AD는 최근 TSAD 분야에서 제안된 혁신적인 방법으로, TimeVQVAE[4] 구조를 바탕으로 TSAD 분야 적용해 성능을 개선한 모델이다. TimeVQVAE-AD는 Masked Generative Modeling을 사용하여 시계열 데이터의 비정상적인 패턴을 탐지하는 기술이다. 이 방법은 시계열 데이터를 시간-주파수 도메인으로 변환하여 다양한 주파수 대역에서 발생할 수 있는 이상을 정교하게 탐지할 수 있도록 설계되었다.

그림 1의 (a)는 Masked Generative Modeling의 과정을 나타낸다. 이 과정에서 시계열 데이터 x 는 STFT(Short-Time Fourier Transform)를 통해 주파수-시간 도메인으로 변환된다. 변환된 데이터는 인코더(Encoder)

를 통해 잠재 표현 z 로 압축된 후, 벡터 양자화(Vector Quantization, VQ) 과정을 통해 고정된 코드북 벡터로 변환된다. 이후, 일부 구간이 마스킹되어 선행 모델(Prior Model)에 입력되며, 이 모델은 마스킹된 데이터를 복원하는 작업을 수행한다. 복원된 데이터는 디코더를 통해 원래의 시계열 데이터 형태로 복원되며, 마지막으로 보간(Interpolation)을 통해 마스킹된 부분이 채워진다. 이상 탐지는 이 복원된 데이터와 원본 데이터 간의 차이를 기반으로 수행된다. 그림 1의 (b)는 TimeVQVAE-AD 모델의 1단계 학습 과정을 설명한다. 이 단계에서는 입력 시계열 데이터가 STFT를 통해 주파수-시간 도메인으로 변환된 후, 인코더를 거쳐 잠재 공간 표현으로 압축된다. 이 잠재 표현은 벡터 양자화 과정을 통해 고정된 코드북 벡터로 변환되며, 이후 디코더를 통해 원래 시계열 데이터로 복원된다. 복원된 데이터는 ISTFT(Inverse STFT)를 통해 다시 시간 도메인 데이터로 변환된다. 그림 1의 (c)는 TimeVQVAE-AD 모델의 2단계인 마스킹과 이상 탐지 과정을 나타낸다. 이 단계에서는 1단계에서 생성된 토큰 일부가 마스킹되며, 마스킹된 데이터는 양방향 트랜스포머(Bidirectional Transformer)에 입력된다. 트랜스포머는 마스킹된 부분을 예측하고, 이를 통해 정상적인 데이터와 비정상적인 데이터를 구분할 수 있는 패턴을 학습한다. 최종적으로, 예측된 토큰은 복원된 데이터로 사용되며, 원래 데이터와의 차이를 바탕으로 이상 점수가 산출된다.

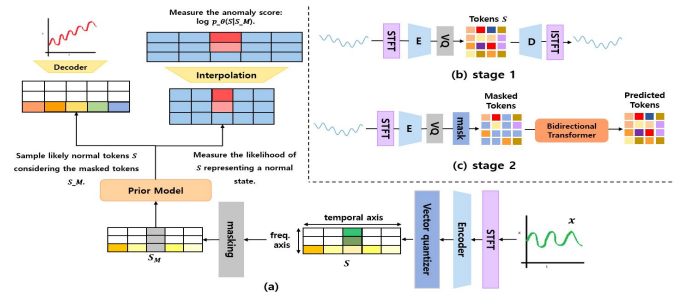


그림 1. Masked Generative Modeling(a)과 Training(b, c) 과정

2.2 설명 가능성(XAI)

설명 가능성은 머신러닝 및 딥러닝 모델에서 점점 더 중요한 요소로 자리 잡고 있다. TimeVQVAE-AD는 이러한 설명 가능성을 제공하기 위해 반사실적(counterfactual) 샘플링을 도입하였다. 반사실적 샘플링은 비정상적인 패턴을 탐지한 후, 그 비정상적인 구간을 제거하고 정상 상태를 재구성하는 과정이다. 이로 인해 사용자는 탐지된 이상이 시스템의 정상 동작과 어떻게 다른지를 직관적으로 이해할 수 있다. 예를 들어, 제조 공정에서 이상이 탐지되었을 때, TimeVQVAE-AD는 그 이상 구간을 정상적인 공정 흐름으로 대체하여, 문제가 발생하지 않았을 경우 시스템이 어떻게 작동했는지를 시각적으로 제공할 수 있다. TimeVQVAE-AD의 설명 가능성은 이러한 요구를 충족시키며, TSAD의 실질적인 활용을 극대화할 수 있는 잠재력을 가진다.

2.3 실험

본 논문에서는 TimeVQVAE-AD의 성능을 검증하기 위해, UCR-TSA 아카이브를 활용한 실험을 설계하였다. UCR-TSA 아카이브는 250개의 고도로 정제된 벤치마크 데이터셋으로 구성되어 있으며, 다양한 시계열 데이터를 포함하고 있어 TSAD 방법론의 성능을 평가하는 데 적합하다. 실험은 TimeVQVAE-AD가 제안한 방식에 따라 수행되며, 각 데이터셋에 대해 탐지된 이상 점수를 분석하고, 이를 기존 연구에서 보고된 결과와 비교한다. 또한, 반사실적 정상 상태 샘플링을 통해 탐지된 이상에 대한 설명 가능성을 평가하고, 이 방법이 실제로 얼마나 유용한지를 검토한다.

2.4 실험 결과 및 분석

그림 2와 그림 3은 UCR-TSA 아카이브의 dataset index 1과 127에 대해 TimeVQVAE-AD와 본 논문에서 독립적으로 수행한 실험 결과를 비교하여 시각적으로 나타낸 것이다. 그림 2의 (a)는 TimeVQVAE-AD를 통해 생성된 결과이며, 그림 2의 (b)는 본 논문에서 동일한 데이터셋에 대해 수행한 실험 결과를 보여준다. 그림의 첫 번째 행은 원본 시계열 데이터를 나타내며, 주황색 선은 TimeVQVAE-AD에 의해 이상으로 탐지된 구간을 강조한다. 두 번째 행에서 'clipped a_s^* '로 표시된 시계열 데이터의 주파수 도메인에서 추출된 이상 점수의 클립(clipped)된 형태를 보여준다. 색상으로 표현된 이 행은 주파수 대역별로 시계열의 비정상적인 활동인 두드러지는 구간을 나타낸다. 세 번째 행은 a_{final} 로 표시된 최종 이상 점수를 시각화한 것이다. 이 점수는 주파수 도메인에서 종합된 이상 점수를 기반으로 하며, 이상 구간에서 급격히 증가하는 모습을 보여준다. 점선으로 표시된 임계값은 이상 여부를 결정하는 기준이 되며, 이를 초과하는 부분이 이상으로 판단된다. 마지막으로 네 번째 행은 반사실적 시계열 데이터를 보여준다. 반사실적 샘플링을 통해, TimeVQVAE-AD는 이상으로 탐지된 구간을 제거하고 정상 상태를 재구성한 결과를 제공한다. 이를 통해 사용자는 비정상적인 구간이 제거되었을 때 시계열이 어떻게 보였는지를 시각적으로 확인할 수 있으며, 이는 제조 공정 등에서 문제 발생 시 원인을 이해하고 정상적인 상태를 파악하는 데 중요한 도구로 작용한다. 실험 결과, TimeVQVAE-AD는 기존 연구에서 보고된 성능을 대체로 재현하며, 다양한 주파수 대역에서 복잡한 이상 패턴을 효과적으로 탐지하는 것으로 나타났다. 특히, 반사실적 샘플링을 통해 제공된 설명 가능성은 사용자가 탐지된 이상을 더 잘 이해할 수 있도록 돕는 중요한 요소로 작용했다. 다만, 일부 데이터셋에서는 기존 연구와의 결과 차이가 발견되었으며, 이는 실험 환경이나 하이퍼파라미터 설정의 차이로 인한 것으로 판단된다. 이러한 차이는 TimeVQVAE-AD의 성능이 특정 조건에서 민감할 수 있음을 시사하며, 추가적인 연구가 필요하다는 점을 보여준다.

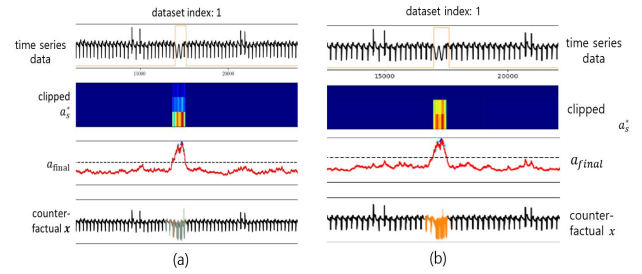


그림 2. UCR-TSA 아카이브 dataset index 1로 실험한 결과, (a)는 TimeVQVAE-AD, (b)는 본 논문에서 실험한 결과

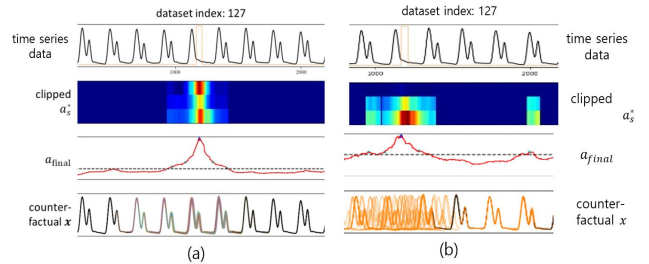


그림 3. UCR-TSA 아카이브 dataset index 127로 실험한 결과, (a)는 TimeVQVAE-AD, (b)는 본 논문에서 실험한 결과

III. 결론

본 논문에서는 TimeVQVAE-AD를 소개하고, 해당 방법론의 성능을 독립적으로 검증하였다. 실험 결과, TimeVQVAE-AD는 높은 탐지 정확도와 우수한 설명 가능성을 제공하며, 대부분의 경우 기존 연구에서 보고된 성능을 재현할 수 있음을 확인하였다. 그러나 일부 데이터셋에서는 성능 차이가 나타나, TimeVQVAE-AD의 실세계 적용 가능성에 대한 추가 검토가 필요함을 확인하였다.

향후 연구에서는 TimeVQVAE-AD의 성능을 더욱 폭넓게 평가하고, 다변량 시계열 데이터에의 적용 가능성을 탐색할 예정이다. 또한, 이 방법론을 다양한 실제 응용 사례에 적용하여 그 유용성을 검증함으로써, TimeVQVAE-AD가 TSAD 분야에서의 새로운 표준으로 자리 잡을 수 있도록 기여하고자 한다.

ACKNOWLEDGMENT

"본 연구는 과학기술정보통신부 및 정보통신기획평가원의 대학ICT연구센터지원사업의 연구결과로 수행되었음" (IITP-2024-2020-0-01808*)

참고 문헌

- [1] A. A. Cook, G. M. S. Irl and Z. Fan, "Anomaly Detection for IoT Time-Series Data: A Survey," in IEEE Internet of Things Journal, vol. 7, no. 7, pp. 6481-6494, July 2020, doi: 10.1109/IJOT.2019.2958185.
- [2] Lee, Daesoo, Sara Malacarne, and Erlend Aune. "Explainable time series anomaly detection using masked latent generative modeling." Pattern Recognition (2024): 110826.
- [3] <https://wu.renjie.im/research/anomaly-benchmarks-are-flawed>
- [4] Lee, Daesoo, Sara Malacarne, and Erlend Aune. "Vector quantized time series generation with a bidirectional prior model." arXiv preprint arXiv:2303.04743 (2023).
- [5] Jae Kwon Bae (2023). A Study on the Applicability of eXplainable Artificial Intelligence(XAI) Methodology by Industrial District. 글로벌경영학회지, 20(2), 195-208.