

A Wavelet-Enhanced Temporal Attention Network for Unsupervised Anomaly Detection in EV Battery Tab-to-Tab Laser Welding

Jeong-Beom Hong*, Byung-Chul Ko**

*Ph.D. course, Dept. of Advanced Materials Processing Engineering, Inha University, Incheon, Korea

**Professor, Advanced Materials & Processing Research Center, Inha University, Incheon, Korea

[Abstract]

Laser welding applies a high-energy-density beam to the material, involving complex physical phenomena, and the measured optical signals exhibit non-stationary characteristics. Defect-induced signal variations appear not only as amplitude changes but also as changes in frequency composition and baseline shifts, and they rarely form fixed patterns corresponding one-to-one with defect types. Accordingly, a model is required that can be trained using only normal data while jointly learning local time-domain patterns and time-frequency characteristics. The proposed model represents both local temporal patterns and time-frequency variations through a hybrid time-frequency embedding combining a multi-scale dilated 1D convolution branch with a Morlet wavelet-based wavelet transform branch, and learns global dependencies among time steps via temporal self-attention. The anomaly detection performance was evaluated through synthetic anomaly injection experiments using various condition combinations and varying anomaly intensities. For actual defects, the model achieved AUROC of 0.925 and 0.941, AUPRC of 0.923 and 0.931, and recall of 0.901 and 0.912 for localized over-welding, corresponding to localized signal variation, and gap defects, corresponding to signal variation over a wide interval, respectively.

▶ **Key words:** Laser welding, Transformer, Wavelet transform, Anomaly Detection, Non-stationary signal, Photodiode signal

[요 약]

레이저 용접은 소재에 높은 에너지 밀도의 빔을 가하기 때문에 복합적인 물리 현상이 수반되며, 측정되는 광학 신호는 비정상성(non-stationary) 특성을 보인다. 결함에 따른 신호 변화는 진폭의 변화뿐 아니라 주파수 구성의 변화와 baseline의 이동으로도 나타나며, 결함 유형과 일대일로 대응되는 고정된 패턴을 이루기 어려운 한계가 존재한다. 이에 정상 데이터만으로 학습하면서 시간 영역의 국부적 패턴과 시간-주파수 특징을 함께 학습할 수 있는 모델이 요구된다. 제안 모델은 Multi-scale dilated 1D convolution branch와 Morlet wavelet 기반 wavelet transform branch를 결합한 hybrid time-frequency embedding을 통해 국소 시간 패턴과 시간-주파수 변화를 함께 표현하고, temporal self-attention을 통해 시점 간 전역적 의존성을 학습한다. 본 연구에서는 모델의 성능 평가를 위해 다양한 조건 조합 및 이상치 세기 변화를 통한 Synthetic Anomaly Injection 실험을 통해 제안 모델의 이상 탐지 성능을 평가하고, 실제 불량 중 국소적 신호 변동에 해당하는 localized over-welding과 넓은 구간의 신호 변동에 해당하는 gap defect에 대해 각각 AUROC 0.925와 0.941, AUPRC 0.923과 0.931, Recall 0.901과 0.912를 기록하였다.

▶ **주제어:** 레이저 용접, 트랜스포머, 웨이블릿 변환, 이상 탐지, 비정상성 신호, 포토다이오드 신호

• First Author: Jeong-Beom Hong, Corresponding Author: Byung-Chul Ko

*Jeong-Beom Hong (boublestep@inha.edu), Dept. of Advanced Materials Processing Engineering, Inha University

**Byung-Chul Ko (bcko87@inha.ac.kr), Advanced Materials & Processing Research Center, Inha University

• Received: 2026. 07. 20, Revised: 2026. 08. 03, Accepted: 2026. 09. 03.

I. Introduction

최근 전 세계적인 탄소 배출 저감 기조에 따른 전기 자동차(Electric Vehicles, EVs) 보급의 급격한 확산은 자동차 제조 산업의 패러다임 혁신을 이끌고 있으며, 이에 따라 핵심 동력원인 고성능 배터리 제조 기술이 시장의 핵심 경쟁력으로 부상하고 있다[1]. 리튬 이온 배터리 팩은 일반적으로 셀-모듈-팩 계층 구조를 따르며, 제품 전체의 성능 및 안정성은 배터리 모듈 내부에 존재하는 수십 개의 전기적 연결부의 접합 품질에 따라 결정된다[2]. 현재 이러한 정밀 접합 기술로 레이저 용접 기술이 활용되고 있으며, 전기 자동차 배터리 외 전장 부품 등 다양한 산업 분야에 핵심적인 비접촉식 접합 기술로 자리 잡고 있다[3].

레이저 용접 공정은 높은 에너지 밀도를 가진 레이저 빔을 이용하여 재료를 용융, 접합한다. 이러한 공정 특성으로 인해 키홀(keyhole) 거동 및 용융 풀(molten pool) 변동, 그리고 플라즈마 방사(plasma emission) 등 물리적으로 복잡한 현상이 수반된다[4]. 이들 현상은 레이저 용접 품질을 판별하는 주요 인자로서, 음향, 광학, CCD 카메라 등을 활용한 센서 기반 monitoring system으로 측정되어 왔다[5].

특히 포토다이오드 기반 모니터링은 용접 시 방출되는 특정 광학 파장 대역을 선택적으로 검출하여 용접 품질을 평가하며, 저렴한 가격 대비 우수한 판별 능력을 갖고 있다. 레이저 용접은 고밀도 에너지를 용접부에 국소적으로 가하여 재료를 급속히 가열·용융시킴으로, 고온 플라즈마의 방사, 키홀 및 용융 풀의 요동, 급격한 열전달과 상변화가 동시다발적으로 발생한다. 이러한 복합적 물리 현상이 중첩된 결과, 포토다이오드로 측정되는 신호는 시간에 따라 주파수 성분과 진폭이 지속적으로 변하고, 일정한 주기 구조를 갖지 않는 비정상성(Non-stationary) 및 비주기성(Aperiodic) 특성을 보인다[6]. 이러한 복잡한 신호 패턴에서 유의미한 정보를 추출하기 위해 하드웨어 필터링과 더불어 디지털 필터링 및 주파수 변환 등이 활용되어왔다. 최근에 딥러닝 시계열 모델 등의 발전으로 인하여 제조 현장의 시계열 신호 패턴을 학습시켜 이상 탐지하는 모델들이 활용되고 있다[7].

그러나 레이저 용접 양산 공정에서 포토다이오드 신호를 이용한 품질 판정과 결함 검출에는 몇 가지 근본적인 어려움이 존재한다. 첫째, 결함에 따른 광학 신호 변화는 결함 유형과 일대일로 대응되는 고정된 패턴으로 나타나지 않는다. 이는 동일한 결함이라도 소재의 물성 및 두께, 접합 상태, 용접 조건 및 외부 온도에 따라 측정되는 신호

변화와 특징을 정량적으로 사전에 규정하기 어려움을 시사한다. 둘째, 양산 현장에서는 결함 데이터가 희소하고 품질 정보가 주로 개별 용접 단위의 합격 또는 불합격 형태로 제공되므로, 결함이 발생한 정확한 시간 구간을 사전에 라벨링 하는 데 한계가 있다.

이러한 어려움으로 인해, 본 연구의 tab-to-tab 겹치기 레이저 용접 환경에서 결함은 검출하기 위해서는 결함 유형별 신호 패턴과 발생 구간을 사전에 명시적으로 정의하기보다 정상 데이터만으로 학습하는 비지도 기반 접근이 요구된다. 또한, 변화는 시간 영역뿐 아니라 주파수 구성에서도 나타날 수 있으므로, 레이저 용접 신호의 국부적 시간 패턴과 시간-주파수 특징을 함께 학습할 수 있는 모델이 필요하다.

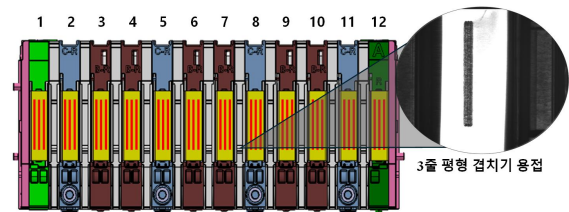


Fig. 1. Battery cell tab-to-tab assembly

이에 본 연구에서는 정상 신호의 변동 구조를 학습하여 그로부터의 이탈을 탐지하는 비지도 재구성 기반 모델을 제안한다. 제안 모델은 특정 결함 유형을 학습하지 않으므로, 탐지 대상은 정상 패턴으로부터 이탈하는 신호 변동의 양상으로 규정하였다. 이러한 관점에서 본 연구는 짧은 구간에 집중되어 나타나는 국소적 신호 변동과 상대적으로 넓은 구간에 걸쳐 지속되는 신호 변동이라는 두 가지 발현 양상을 대상으로 진행된다.

두 양상은 각각 양산 공정에서 라벨 확보가 가능한 국소 과용접(localized over-welding)과 갭 불량(gap defect)에 대응한다. 국소 과용접은 좁은 간격의 3중 연속 용접에서 인접 비드의 잔열이 축적되는 누적 입열 현상에 의해 국소적으로 발생하며, 짧은 구간에 집중된 신호 증폭과 신호의 baseline 상승이 특징이다.

갭 불량은 상·하부 모재 사이의 간극에 의해 계면 열전달이 제한되어, 상대적으로 넓은 구간에 걸쳐 신호 변동의 심화를 동반한다. 다만 이 동일한 결함 유형이라도 발생 조건에 따라 신호 발현 형태가 달라질 수 있어 두 유형이 항상 동일한 신호 패턴을 보인다고 일반화할 수는 없으나, 제안 모델은 결함 유형별 패턴이 아닌 정상 패턴으로부터의 이탈에 반응하므로 이러한 발현 형태의 다양성은 학습 과정에 제약으로 작용하지 않는다.

본 논문의 구성은 다음과 같다. II장에서는 레이저 용접 다변량 신호와 딥러닝을 활용한 이상 탐지에 관한 관련 연구를 정리한다. III장에서는 본 연구의 레이저 용접 데이터 취득 시스템 설명과 전처리를 설명한다. IV장에서는 제안된 모델의 구조 및 학습 환경 그리고 Loss function, Anomaly scoring에 대해 기술 한다. V장에서는 현장 레이저 용접 불량에 대한 비지도 이상 탐지 결과를 분석한다. VI장에서는 본 연구의 결론과 향후 연구 방향을 제시한다.

II. Preliminaries

2.1 Deep Learning Approaches for Welding Monitoring

레이저 용접 공정에서는 금속 재료의 용융 및 응고뿐만 아니라 기공과 용융 풀의 거동, 플라스마 방사 등 복합적인 물리 현상이 수반되며, 이로부터 다양한 공정 신호가 발생한다. 이에 따라 여러 센서에서 취득한 정보를 활용하는 다중 센서 기반 용접 모니터링 연구가 활발히 수행되어 왔다[8][9].

최근에는 딥러닝 기술의 발전과 함께 다중 센서로부터 취득한 공정 신호를 학습하여 용접 품질을 예측하거나 결함 유형을 분류하는 연구가 증가하고 있다[10]. 특히 포토다이오드 신호를 활용한 용접 품질 판정 및 이상 탐지 연구도 지속적으로 수행되고 있다. 예를 들어, Niu 등[11]은 원시 포토다이오드 신호에 웨이블릿 패킷 변환(Wavelet Packet Transform, WPT)을 사전 적용하여 주파수 대역별 특징을 추출한 후, 이를 CNN 기반 모델에 입력하여 레이저 용접의 용입 상태를 분류하는 지도학습 방법을 제안하였다. 또한, Zhang 등[12]은 포토다이오드 신호와 고속 카메라 영상을 함께 활용한 지도학습 기반 CNN 모델을 제안하였다. 이 연구에서는 포토다이오드 신호를 웨이블릿 패킷 분해(Wavelet Packet Decomposition)를 통해 16개의 주파수 대역으로 분해하여 시간-주파수 영역 특징을 추출하고 이를 CNN 모델 입력으로 활용하여 모델의 성능을 향상시켰다. Ozkat [13]은 레이저 용접의 비정상성 신호 특성을 Wavelet Scattering과 1D convolution Auto-Encoder를 활용한 비지도 학습 이상 탐지 모델을 제안했으며, 기존 지도학습 모델과 비교를 통해 우수한 정확도를 보였다. 이전의 선행연구들은 다중 센서 또는 포토다이오드 신호에 딥러닝을 적용하여 레이저 용접 품질 판정과 결함 탐지의 가능성을 보여주었다. 그러나 양산 제조 현장에서는 결함 데이터가 희

소하고, 특정 결함 발생 시점에 대한 세밀한 라벨을 확보하기 어렵다. 따라서 결함 유형과 발생 구간에 대한 사전 라벨 없이 정상 데이터만을 이용하여 학습할 수 있는 비지도 이상 탐지 접근이 요구된다. 또한, 앞서 검토한 웨이블릿 기반 연구에서는 시간-주파수 특징이 모델 학습 이전의 별도 변환 단계에서 생성되었다. 이와 달리 사전 시간-주파수 변환에 의존하지 않고, 다변량 원본 신호로부터 특징 표현 학습과 재구성까지 End-to-End로 학습하는 모델 연구는 레이저 용접 이상 탐지 분야에서 여전히 필요하다.

이에 본 연구는 이러한 한계를 보완하기 위해, 다변량 광신호와 레이저 출력 신호를 원시 형태로 입력받아 특징 추출부터 재구성까지 End-to-End로 학습하는 비지도 이상 탐지 모델을 제안한다. 제안 모델은 Transformer 기반의 비지도 재구성 구조를 활용하며, 배터리 팩 모듈 조립 양산 공정의 tab-to-tab 겹치기 용접에서의 정상 데이터로만 학습하여, 이상 신호를 탐지한다.

III. Experimental setup

3.1 Signal Acquisition

본 연구는 전기자동차 배터리 모듈 조립 공정(Battery Module Assembly, BMA)에서 배터리 셀 전극과 탭 간의 겹치기 레이저 용접을 대상으로 하며, 하나의 모듈당 총 36회의 tab-to-tab 겹치기 용접이 순차적으로 진행되며, 각 용접부는 좁은 간격(0.1mm)으로 배치된 3줄의 용접이 진행된다. 재료는 상부 모재로 알루미늄(A1050-H0), 하부 모재로 구리(C1100-1/2H)를 사용하였다. 용접에는 TRUMPF사의 TruDisk 3001 디스크 레이저 시스템을 사용하였다.

본 연구에서는 레이저 용접 공정 중 발생하는 다채널 신호를 취득하기 위해 Fig. 2와 같은 포토다이오드 기반 레이저 용접 monitoring system을 구성하였다.

취득 신호는 근적외선(Near-Infrared, NIR) 및 가시광선(Visible-light, VIS) 두 대역을 검출하는 두 개의 포토다이오드 채널과 디스크 레이저에서 제공되는 하나의 레이저 입출력 (Laser output, L/O) 제어 채널로 구성된다. 용접부에서 방출된 광은 동축 광학계를 통해 수집된 후, Beam splitter에 의해 두 개의 광 경로로 분기된다.

첫 번째 경로는 long-pass filter(통과 대역: 1,100~1,700nm)를 거쳐 근적외선(Near-Infrared, NIR) 포토다이오드로, 두 번째 경로는 band-pass filter(통과 대역: 500~900nm)를 거쳐 가시광선(Visible-light, VIS) 포토다이오드로 각각 검출된다.

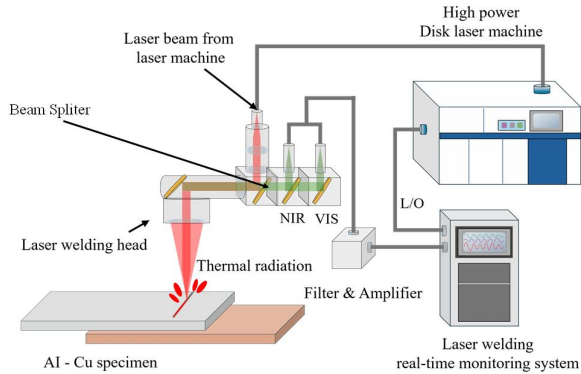


Fig. 2. The multi-sensor laser welding system

두 광학 채널 중 NIR은 용접부의 전반적인 열적 상태를 반영하며, VIS는 상부 모재의 용융 풀 표면에서 방사되는 열복사의 단파장 성분과 일부 플라스마 방출 정보를 포함한다. 레이저 입/출력 제어 채널(Laser-output, L/O)은 Disk 레이저 기기에서 출력되는 아날로그 전압 신호로서, 공정 중 레이저 출력 수준의 변화를 0-4V 범위의 전압으로 반영한다.

세 채널의 아날로그 전압 신호는 DAQ 장치를 통해 샘플링 주파수(f_s) 10.72kHz로 디지털 변환되어 기록되며, 이때 NIR, VIS 신호 값은 절대적인 광량이 아닌 광 강도에 비례하는 상대적 수치로 취급된다. 이에 따라 원본 데이터는 시점당 3개 변수(NIR, VIS, L/O)로 구성된 다변량 시계열을 형성하며, 개별 용접 신호는 약 1,080 시점 내외의 길이를 갖는다.

Table 1. Specifications of the acquired signals

Channel	Detection band / Range	Measured quantity	Signal type
NIR (Near-Infrared)	1,100-1,700 nm	Weld thermal radiation	Photodiode voltage (a.u.)
VIS (Visible-Light)	500-900 nm	Molten pool, Plasma emission	Photodiode voltage (a.u.)
L/O (Laser output)	0-4 V	Laser power level	Analog voltage

3.2 Signal data and pre-processing

본 연구는 BMA 공정에서 수집한 200개의 배터리 모듈 생산 데이터를 학습 및 평가에 활용하였다. 하나의 배터리 모듈 생산에는 총 36회의 tab-to-tab 겹치기 용접이 순차적으로 진행되며, Fig. 1에서 보여주고 있다.

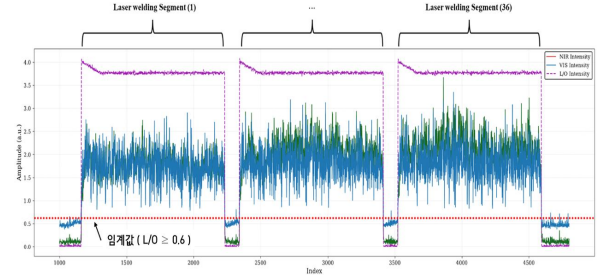


Fig. 3. Representative measured waveforms of NIR, VIS and L/O signals during tab-to-tab laser welding

모듈 단위로 저장된 연속 신호를 개별 용접 단위로 구분하기 위해 segmentation을 수행하였다. 레이저 조사 구간을 식별하기 위해 L/O 신호를 기준으로 하였으며, L/O 신호가 0.6V 이상으로 상승하는 시점을 레이저 조사 시작점으로, 다시 0.6V 미만으로 하강하는 시점을 조사 종료점으로 정의하였다. 두 지점 사이의 연속 신호 구간을 하나의 개별 용접 segment로 추출하였다. Fig. 3은 측정된 원본 레이저 용접 다변량 신호를 보여준다.

정상 용접 데이터는 총 7,200개의 개별 용접 segment로 구성되며, 70:15:15 비율을 통해 Train 데이터는 5,040개와 Validation, Calibration 데이터는 1,080개 데이터로 분할 하였다. Validation 데이터는 모델 파라미터 업데이트에는 사용하지 않고, 각 epoch 종료 후 재구성 손실을 산출하여 학습을 스케줄러의 갱신과 조기 종료(early stopping) 여부를 판단하는 데 사용하였다.

3.3 Synthetic Anomaly Injection

본 절에서는 현장에서 측정된 정상 레이저 용접 신호에 대한 합성 이상 주입 방법을 기술한다.

앞서 기술한 바와 같이 본 연구의 품질 정보는 작업자의 전수 표면 검사를 통해 개별 용접 단위로 합격 또는 불합격이 판별된 형태이며, 결함이 발생한 시간 구간은 사전에 라벨링 되어 있지 않다. 따라서 이상이 발생한 시점과 지속 구간을 알 수 없어, 모델이 이상을 어떤 조건에서 얼마나 민감하게 포착하는지를 실제 결함 데이터만으로 분석하기 어렵다.

이에 본 연구에서는 이상의 발생 위치, 지속 길이, 세기를 통제할 수 있는 합성 이상 주입을 도입하여, 사전 라벨링을 하지 못한 상황에서 모델의 성능을 평가한다. 합성 이상 신호는 용접 신호의 특징을 반영하기 위해, 용접부의 평균적 상태를 반영하는 DC 오프셋 성분(Baseline)과 킥홀 진동 및 용융 풀 요동 등 동적 물리 현상을 반영하는 AC 변동 성분의 중첩으로 구성된다. 합성 이상은 이 두 성분을 각각 조작하여 생성하였다.

식 (1)은 레이저 용접 신호 특성에 맞춰진 합성 이상 주입 수식이다.

$$x_{anomal}(t) = x^{(c)}(t) + w(t)[(\alpha_{target} - 1)(x^{(c)}(t) - \mu_{normal}) + DC_{shift} + \beta \cdot n[f_1, f_2](t)] \quad (1)$$

식(1)의 $x^{(c)}(t)$ 는 광학 채널 $c \in \{NIR, VIS\}$ 의 원본 정상 신호이며, μ_{normal} 은 이상 주입 구간 이전 시점에서의 신호 평균이다.

$(\alpha_{target} - 1)(x^{(c)}(t) - \mu_{normal})$ 은 AC 변동 성분에 대한 향으로, 정상 신호의 평균 대비 변동 성분을 α_{target} 을 통해 조정한다.

DC_{shift} 는 DC 오프셋 성분에 대한 향으로, 해당 값이 커지면 주입 구간 내 baseline의 상승을 표현한다. 추가로 $\beta \cdot n[f_1, f_2](t)$ 항은 특정 주파수 성분이 증가하기 위한 향으로 특정한 주파수 대역의 잡음을 추가한다. 본 연구에서는 $f_1 = 1,500Hz$, $f_2 = 3,500Hz$ 대역에서 생성되며 진폭 계수 β 는 0.2로 고정하였다.

$w(t)$ 는 주입 구간의 가중 함수로, 식 (2)와 같이 상승 항 $A_{on}(t)$ 와 종료 항 $A_{off}(t)$ 의 곱으로 정의된다. 여기서 t_0 와 t_1 는 각각 이상 주입 구간의 시작과 종료 시점이며, τ 는 전이 속도를 결정하는 시정수이다. 본 연구에서는 $\tau = 1$ 로 설정하였는데, 이는 전이 구간이 길어질 경우 이상이 실제 발현되는 구간과 사전에 정의한 주입 구간 $[t_0, t_1]$ 사이에 불일치가 발생하여 sliding window 기반 anomaly scoring의 평가 정확성이 저하되기 때문이다.

$$A_{on}(t) = \begin{cases} 0, & t < t_0 \\ 1 - \exp(-\frac{t-t_0}{\tau}), & t \geq t_0 \end{cases} \quad (2)$$

$$A_{off}(t) = \begin{cases} 1, & t < t_1 \\ \exp(-\frac{t-t_1}{\tau}), & t \geq t_1 \end{cases}$$

Table 2는 이상 주입의 지속 길이와 발생 위치를 조합한 9가지 시나리오이며, 지속 길이 3수준(Short: 150, Mid: 250, Long: 300)과 발생 위치 3수준(Early, Center, Late)으로 구성된다. 두 축은 각각 다음의 근거에 따라 설정하였다. 지속 길이 축은 서론에서 기술한 두 가지 결함 발현 양상, 즉 짧은 구간에 집중되는 국소적 신호 변동과 넓은 구간에 걸쳐 지속되는 신호 변동을 함께 포괄하기 위한 것이다. 현장 결함의 신호 발현 형태는 사전에 정량적

으로 규정하기 어려우므로, 평가가 특정 발현 길이에 대한 가정에 종속되지 않도록 짧은 이상부터 긴 이상까지 세 수준으로 구성하였다.

발생 위치 축은 서론에서 기술한 것과 같이 국소적 과용접은 특정한 위치보다는 3줄 연속 용접을 통해 누적 입열에 의한 과용접이기 때문에, 발현 위치가 매우 확률적이기 때문에, 모델이 발생 위치에 무관하게 이상을 포착하는지를 확인할 필요가 있으며, 이를 위해 용접 구간의 초반, 중반, 후반에 각각 이상을 주입하였다.

Table 3은 합성 이상의 세기를 5가지 레벨로 정의한다. 두 조정 변수는 식 (1)의 신호 성분에 각각 대응하며, α_{target} 은 AC 변동 성분의 증폭 배율을, DC_{shift} 는 주입 구간 내 baseline의 상승량을 조정한다. 국소적 과잉 입열 시 광신호의 변동 진폭과 baseline이 함께 상승하는 발현 특성을 반영하기 위해, 두 변수를 점차 증가하도록 배치하였다.

5장에서는 위 9가지 시나리오와 5가지 세기 레벨을 조합한 총 45가지 주입 조건에서 평균 AUROC와 AUPRC를 산출하여 비교 모델과의 성능 평가에 활용한다. 비교 모델로는 기존 시계열 재구성 기반 모델과의 비교로, 다변량 시계열을 각 시점의 벡터로 처리하는 RNN (Recurrent Neural Network) 기반 Auto-Encoder (RNN-AE)와 LSTM (Long Short-Term Memory) 기반 Auto-Encoder(LSTM-AE)를 대상으로 한다.

또한, Table 3의 이상 세기를 Level 2로 고정한 상태에서 9가지 시나리오를 통해 4.3절 Anomaly scoring의 핵심 파라미터인 sliding window size(W)와 top-K 비율의 도출에 활용하며, 결과는 5장에서 보여준다. 그리고 생성된 합성 이상이 실제 결함(Localized Over-welding, Gap Defect)에서 관측되는 신호 변동 수준을 포괄하는지는, 5장에서 정량적 비교를 통해 검증한다. Fig. 4는 주입 구간에 따라 합성 이상이 다르게 적용되는 예를 보여준다.

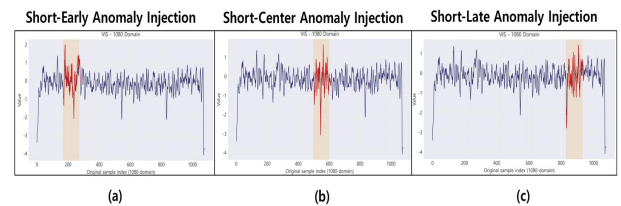


Fig. 4. Examples of synthetic anomaly injection in laser welding signals

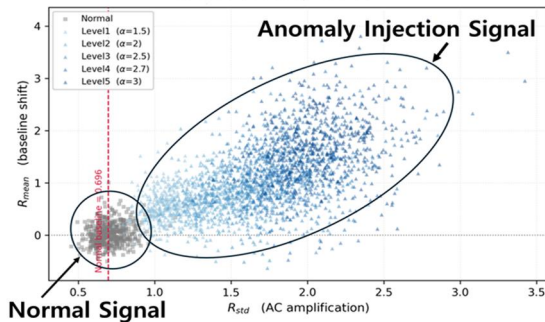
Table 2. Synthetic Anomaly Injection scenarios defined by anomaly duration and injection location.

ID	Duration	Position	t_0	t_1
S01	Short(150)	Early	170	320
S02	Short(150)	Center	500	650
S03	Short(150)	Late	780	930
S04	Mid(250)	Early	170	420
S05	Mid(250)	Center	500	750
S06	Mid(250)	Late	680	930
S07	Long(300)	Early	170	470
S08	Long(300)	Center	500	800
S09	Long(300)	Late	630	930

Table 3. Severity levels for synthetic anomaly injection.

Level	α_{target}	DC_{shift}
Level 5	3.0	1.0
Level 4	2.7	0.7
Level 3	2.5	0.5
Level 2	2.0	0.4
Level 1	1.5	0.3

(a) Distribution of Normal signals and Synthetic anomalies at five severity levels



(b) Real defects overlaid on Synthetic Anomaly severity levels

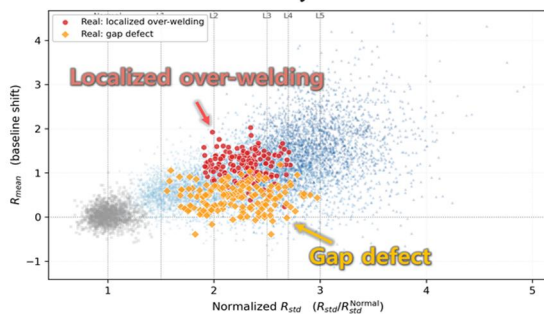


Fig. 5. Signal deviation levels of synthetic anomalies and real defects: (a) normal signals and five synthetic severity levels; (b) real defects overlaid on the synthetic levels.

본 절에서 생성한 합성 이상과 서론에 언급한 실제 결함 (Localized over-weld, Gap defect) 간 분포 유사성 검증이 필요하다. Fig. 5(a)는 정상 신호와 Table 3의 이상 신호 세기 범위 (Level 1~5)의 R_{std} , R_{mean} 분포를 보여준다. 유사성 검증을 하기 위해, 이상 구간 $[t_0, t_1]$ 내부의

표준편차와 평균을 s_{in} , m_{in} 계산하고, 그 외 구간의 표준편차와 평균을 s_{out} , m_{out} 계산하였다. AC 변동 성분의 증

폭 정도를 나타내는 $R_{std} = \frac{s_{in}}{s_{out}}$ 수식으로 계산되며,

baseline 이동량을 나타내는 $R_{mean} = \frac{(m_{in} - m_{out})}{s_{out}}$ 수

식으로 도출하였다. Fig. 5(b)는 실제 결함을 이상 신호 세기 범위에 오버랩한 결과를 보여준다. 정규화된 R_{std} 을 기준으로 국소 과용접은 약 2.0~2.8, 갭 불량은 1.9~2.8 범위에 분포하고 있다. 두 결함 모두 합성 이상의 Level 2~4구간에 해당하는 신호 변동 수준을 보였다.

이는 Table 2에서 설정한 세기 범위가 실제 결함에서 관측되는 변동 수준을 중심에 두고, Level 1(약한 이상)과 Level(강한 이상)까지 포함하도록 구성되었음을 의미한다.

또한, Fig. 5를 통해 알 수 있는 것은 두 가지 불량으로 구분하였으나, 실제 결함에 대한 신호 편차는 매우 다양하고, 사전 라벨링 하기 어렵다는 것을 간접적으로 알 수 있다.

그리고 R_{mean} 축에서는 두 결함이 구분되는 양상이 나타나는데, 국소 과용접은 갭 불량에 비해 높은 baseline 상승을 동반한다는 것을 알 수 있다. 반면, 갭 불량은 baseline 변화보다 변동 진폭의 증가가 주된 발현 양상임을 확인하였다.

IV. Proposed Method

본 연구는 레이저 용접 공정에서 취득된 다변량 포토다이오드 신호와 레이저 출력 신호의 정상 패턴을 학습하여 이상을 탐지하는 비지도 재구성 기반 모델을 제안한다. 레이저 용접 신호는 keyhole 거동, 용융 풀 요동, plasma 방사 등 복합적인 물리 현상이 중첩되어 매우 비정상성과 비주기성을 갖는 신호 패턴을 보인다. 이러한 특성은 재구성 기반 이상 탐지에서 정상 공정 변동과 결함에 의한 이상 변동을 구분하기 어렵게 만든다.

이에 본 연구에서는 Wavelet-enhanced Temporal Attention(WTA) 모델을 제안한다. 제안 모델은 시간 영역의 다중 해상도 국소 패턴과 시간-주파수 영역의 특징을 병렬로 추출하는 hybrid embedding 구조를 통해 각 시점의 표현을 구성하고, 이를 기반으로 temporal attention이 시점 간 전역적 의존성을 학습하도록 설계되었다. 본 장에서는 embedding 구조(4.1절), temporal attention(4.2절), 손실 함수 및 anomaly score(4.3절)를 순서대로 기술한다.

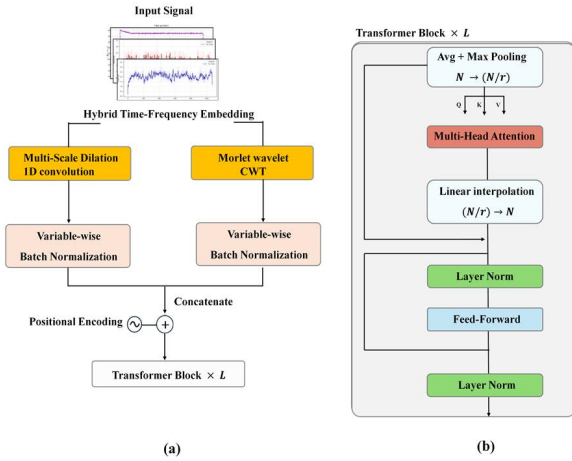


Fig. 6. Overall architecture of the proposed model: (a) Embedding structure (b) Transformer block

4.1 Hybrid time-frequency embedding

Transformer의 self-attention은 시계열의 모든 시점 간 전역적 의존성을 효과적으로 모델링할 수 있다는 장점을 갖는다. 다만 self-attention은 입력 token 간 Query와 Key의 유사도에 기반하여 Value를 가중합하는 구조이므로, attention이 포착하는 의존성의 품질은 사전에 구성된 token 표현의 품질에 크게 좌우된다[14].

이에 제안 모델은 Fig. 6(a)과 같이 서로 다른 두 가지 embedding 방식을 병렬로 적용하여, 레이저 용접 신호의 비정상성 및 비주기적 특성을 표현한다.

첫 번째 multi-scale convolution branch는 서로 다른 커널 크기(kernel size, K)와 팽창 계수(dilation factor, D)를 갖는 J 개의 non-causal dilated 1D convolution layer를 병렬로 적용한다.

Fig. 7은 각 layer의 커널 크기, 팽창 계수, 출력 채널 수를 보여준다. 본 연구에서 병렬 layer의 수 J 는 3으로 설정 하였으며, 각 layer의 (K, D) 조합은 (4, 1), (8, 2), (16, 4)로 구성하였다. Dilated convolution의 receptive field는 $(K-1) \times D + 1$ 로 산출되므로 각 layer는 4, 15, 61 시점의 receptive field를 가지며, 이는 샘플링 주파수 10.72kHz 기준으로 약 0.37ms, 1.4ms, 5.7ms의 시간 구간에 해당한다. 이와 같은 다양한 receptive field를 통해 순간적 신호 변화와 상대적으로 완만한 신호 변화를 포착하며, 각 layer 출력은 42+43+43으로 128채널의 시간 영역 표현을 형성한다.

두 번째는 레이저 용접 과정에서 발생하는 포토다이오드 신호에서는 keyhole 진동 (약 500Hz ~ 2kHz), 용융 풀 (약 50 ~ 500Hz), spatter 발생 (약 2 ~ 5kHz) 등 서로 다른 물리적 현상에 기인한 주파수 대역이 공존한다. 이러한

복합적인 신호 특성으로 인해, 기존 연구에서는 시간 정보를 주파수 정보로 변환해 왔다[15].

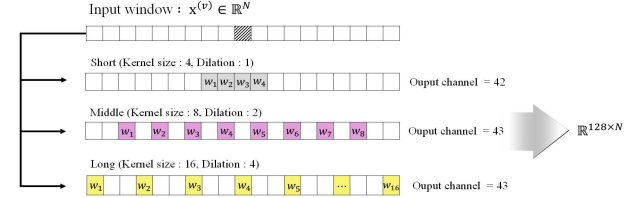


Fig. 7. Configuration of Non-causal dilated 1D convolution

Continuous wavelet transform (CWT)는 이전 Fast Fourier transform, Short-time Fourier transform과 다르게, 고정된 크기의 window를 사용하지 않고, 주파수 대역에 따라 scale을 가변적으로 조절하여 신호를 분석하는 장점이 있다. 이에 본 연구는 CWT branch로 wavelet transform을 활용하며, mother wavelet 함수로 Morlet wavelet을 채택하였다.

Morlet wavelet은 Gaussian envelope와 복소 지수함수의 곱으로 정의되며, 시간 및 주파수 영역 모두에서 우수한 국소화 특성을 보인다. 식 (3)은 Morlet wavelet의 수식을 나타내며, 여기서 ω_0 는 wavelet의 형상을 결정하는 중심 각 주파수로 6.0을 사용하였다[16].

$$\psi(\eta) = \pi^{-1/4} e^{i\omega_0\eta} e^{-\eta^2/2} \quad (3)$$

앞서 기술한 바와 같이 CWT는 고정된 크기의 window 대신 scale(s)을 조절하여 신호를 분석하며, 이때 scale(s)는 mother wavelet을 시간 축으로 늘리는 정도를 나타내는 무차원 파라미터로, 각 scale에 대응하는 분석 주파수는 식 (4)과 같이 산출된다.

$$f = \frac{\omega_0}{2\pi \cdot s \cdot \Delta t}, \Delta t = \frac{1}{f_s} \quad (4)$$

식 (4)에서 알 수 있듯이 분석 주파수는 scale(s)에 반비례하므로 s 가 작을수록 좁은 wavelet이 고주파 성분을, 클수록 넓은 wavelet이 저주파 성분을 분석하게 된다. 이를 통해 레이저 용접 포토다이오드 신호의 고주파 변화와 완만한 주파수 변화를 CWT branch에서 특징으로 활용한다. 본 연구에서는 레이저 용접 과정에서 나타나는 물리 현상의 주파수 대역을 포괄하도록 scale(s)의 범위를 $2 \leq s \leq 100$ 으로 설정하였다. $f_s = 10,720 \text{ Hz}$ 일 때, s

= 100은 약 102Hz, $s=2$ 는 약 5,118Hz에 대응하므로 분석 주파수 범위는 약 102Hz ~ 5,118Hz에 해당하며, 이는 앞서 언급한 용융 풀, Keyhole 진동 및 Spatter 발생의 주파수 대역을 포함한다. 또한, 이 범위의 상한은 샘플링 주파수의 Nyquist 한계 ($f_s/2 = 5,360 \text{ Hz}$) 이내로, 모든 scale의 분석 주파수가 취득 신호에서 유효하게 복원 가능한 대역에 해당한다. 이에 $\text{scale}(s)$ 범위를 [2,100]으로 설정하고, 이 구간을 로그 간격으로 이산화하여 사용한다. S 는 이산화된 $\text{scale}(s)$ 값의 개수를 나타내며, 본 연구에서는 128개의 이산화 값을 활용한다.

이렇게 정의된 scale 집합은 $S=\{s_1, s_2, \dots, s_S\}$ 표현되며, $s_1=2, s_S=100$ 의 범위를 갖고 식(4)에 대입하여 해당 $\text{scale}(s)$ 의 분석 주파수(f)로 환산되어 입력 신호의 시간-주파수 표현을 계산한다. 그 결과 각 변수 (v)는 $N=1,080$ 의 시계열 길이를 갖고 있기에, CWT 결과는 $S \times N$ 크기의 magnitude map을 구성한다. 해당 map의 각 원소는 대응하는 scale과 시점에서의 국소 주파수 성분 강도를 나타낸다.

Fig. 8에서는 각 변수의 $S \times N$ magnitude map이 두 개의 Convolution 2D layer를 통과하는 것을 보여준다. 첫 번째 layer는 인접한 scale과 시점 영역에서 나타나는 저수준 국소 time-scale 패턴을 16개의 feature map으로 추출하고, 두 번째 layer는 이 16개의 feature map을 조합하여 128개의 time-scale feature map으로 변환한다. 이후 adaptive average pooling을 통해 scale 축만을 요약하고, 시간 축(N)은 원래 해상도를 유지한다. 이는 시간 축으로 이후 Transformer block의 temporal attention이 시점 단위로 동작하며, 이상이 발생하는 시점을 국소적으로 포착하기 위함이다. 이에 따라 CWT branch는 각 변수에 대해 $(N, C_{\text{CWT}})=(1080, 128)$ 크기의 시간-주파수 표현을 출력한다.

최종적으로 두 branch의 feature map을 채널 차원에서 결합하여 256 크기의 hybrid representation을 구성하고, linear projection을 통해 최종 embedding 벡터로 변환한다. 이를 통해 최종 embedding 텐서는 $(B, V, N, d_{\text{model}})$ 형태로 형성되어 후속 Temporal Attention 입력으로 사용된다.

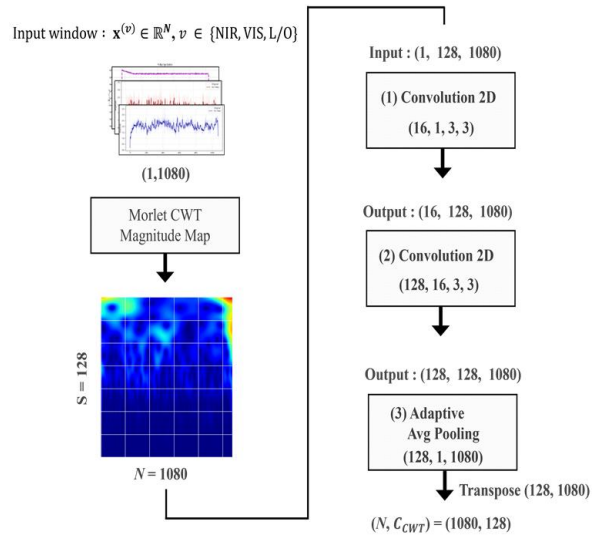


Fig. 8. CWT Branch

4.2 Temporal Attention

본 모델의 입력 신호 길이는 $N=1080$ 이며, 이는 temporal self-attention의 계산 복잡도가 $O(N^2)$ 이므로 입력 길이의 제곱에 비례하여 증가한다. 이는 학습 시 메모리 부담을 일부 초래한다. 이를 완화하기 위해, attention 연산 이전에 시간 축을 Down-sampling 하여 계산 효율성을 확보한다. 이를 위해 Average pooling과 Max pooling을 각각 수행한 뒤, 두 결과를 평균하여 길이가 N/r 으로 축소된 신호로 구성된다. 제안된 모델에서는 Down-sampling(r) 비율을 2로 설정 하였다.

Down-sampling 된 이후 신호 $X_{\text{down}}^{(v)}$ 는 시점 간의 관련도를 산출하기 위한 Query, Key, Value 행렬을 생성한다. 식 (5)는 세 행렬이 학습 가능한 선형 사상을 통해 생성되는 과정을 나타낸다. 여기서 W_Q, W_K, W_V 은 가중치 행렬이며, 그 결과로 얻어진 $Q^{(v)}, K^{(v)}, V^{(v)}$ 는 각각 temporal attention의 Query, Key, Value 행렬을 의미한다. Attention 연산 후에는 선형 보간을 통해 원래 길이 N 으로 복원하고, 잔차 연결과 레이어 정규화를 적용한다.

위 설계를 통해 attention의 계산 복잡도는 $O(N^2)$ 에서 $O((N/r)^2)$ 로 감소하며, 이는 메모리의 연산 효율성을 높였다. Fig. 6(b)는 Transformer block의 전체적인 흐름을 보여준다. 본 연구에서는 Transformer block의 수를 3으로 설정하였으며, 이는 시점 간 의존성을 계층적으로 학습하도록 하였다.

$$\begin{aligned} Q^{(v)} &= X_{down}^{(v)} W_Q, \\ K^{(v)} &= X_{down}^{(v)} W_K, \\ V^{(v)} &= X_{down}^{(v)} W_V \end{aligned} \quad (5)$$

$$Attn^{(v)} = \text{Softmax}\left(\frac{Q^{(v)} K^{(v)T}}{\sqrt{d_h}}\right) V^{(v)} \quad (6)$$

Table 4. Architecture summary of the proposed model

Stage	Component	Configuration	Output Shape	Params
Embedding	Multi-scale convolution branch	$J=3,$ $(K,D)=(4,1),(8,1),(16,4)$	$(V, N, 128)$	1,328
	Morlet CWT branch	$S=128,$ $\omega_0=6.0$	(V, S, N)	0
	Conv2D + AdaptiveAvg pool	Featuremap $1 \rightarrow 16 \rightarrow 128$	$(V, S, 128)$	18,720
	Concat + Linear projection	$256 \rightarrow d_{model}$	$(V, S, 256)$	65,792
Encoder	Temporal Transformer Block $\times L$	$L=3,$ Multi-head = 4, Down-sampling = 2 $FFN=4$	$(V, N, 256)$	2,366,208
Output	Reconstruction head	$d_{model} \rightarrow 1$	$(V, N, 1)$	33,025
Total				~ 2.49M

제안 모델의 전체 구조는 Table 4와 같다. Embedding 단계에서 각 변수는 multi-scale convolution branch와 CWT branch를 병렬로 통과하여, 각각 128개 채널의 표현을 생성하고, 이후 결합된 256차원 표현이 linear projection을 거쳐 d_{model} 차원의 embedding으로 변환된다. 이후 3개의 temporal transformer block을 통과하여 재구성 신호를 출력한다. 제안 모델의 학습 가능 파라미터는 총 약 2.49 M이다.

4.3 Loss function & Anomaly Score

제안 모델은 정상 데이터에 대한 재구성 오차 (reconstruction error)를 기반으로 이상 여부를 판별한다. 본 절에서는 학습에 사용되는 손실 함수와 재구성 오차로부터 anomaly score를 산출하는 과정을 기술한다.

모델에 입력되는 개별 용접 신호 $X \in \mathbb{R}^{N \times V}$ 와 모델의 재구성 출력 $\hat{X} \in \mathbb{R}^{N \times V}$ 의 차이는 평균 제곱 오차(Mean

Squared Error, MSE)를 통해 산출한다.

제안하는 anomaly scoring은 sliding window top- K mean aggregation 방법을 활용하며, 절차는 다음과 같다.

(1) 입력 신호와 재구성 신호 간의 시점별·변수별 reconstruction error를 계산한다.

(2) 개별 용접(세그먼트) 단위의 신호를 sliding window로 분할하고, 각 window 내에서 앞서 계산된 error값 중 상위 $K\%$ 에 해당하는 값들을 평균한다. 이 값은 해당 window와 해당 채널을 대표하는 score로 정의한다.

(3) 채널별로 산출된 score는 각 신호의 물리적 역할에 따라 그룹화한다. NIR과 VIS는 동일한 용접 현상을 서로 다른 파장 대역에서 상보적으로 관측하므로, 어느 한 채널에서 나타난 이상 반응이 다른 채널 값에 의해 희석되지 않도록 $S_{NV} = \max(S_{NIR} + S_{VIS})$ 집계한다. 반면 L/O는 레이저 출력 수준을 반영하는 공정 입력 신호로서 단조로운 신호 특성을 갖는다. 이에 하나의 전역 score로 통합할 경우 신호별 동적 범위와 재구성 난이도의 차이로 인하여 특정 채널의 이상 반응이 희석되거나 과대 반영되는 문제를 방지하기 위해 분할 계산한다.

모델 성능 평가 및 hyper-parameter 도출에는 광학 채널을 기준으로 평가하였으며, 각 $AUROC_{NV}, AUPRC_{NV}, \tau_{NV}$ 에 해당한다.

V. Experimental Results

5.1 Implementation Details

본 절에서는 모델의 학습 환경 및 학습 설정을 기술한다. 제안 모델은 Pytorch(v2.1.0, Python 3.12, CUDA 12.8) 환경에서 구현하였으며, 학습은 Google Colab의 NVIDIA A100-SXM4-40GB GPU에서 수행하였다.

최적화 알고리즘으로 Adam optimize를 사용하였고, 학습 초기의 안정성을 확보하기 위해 warmup 스케줄러를 적용하였다. 매 epoch마다 validation 데이터셋에 대한 재구성 손실을 산출하여, 해당 손실이 일정 횟수 이상 개선되지 않는 경우 학습을 조기 종료하고, 최저 validation loss를 기록한 시점의 모델을 최종 모델로 채택하였다.

비교 모델인 RNN과 LSTM은 기존 시계열 패턴 학습기반에서 사용되는 대표적인 알고리즘으로 알려져 있다[17].

제안 모델인 WTA와 비교하기 위해 Auto-encoder를 결합한 비지도 재구성 기반 이상 탐지 모델로 설계하였다. 자세한 비교 모델에 대한 설정은 Table 5에 정리하였다.

RNN-AE와 LSTM-AE의 encoder 및 decoder는 동일한 대칭형 구조로 구성하였다. Encoder는 두 개의 recurrent layer로 구성되며, hidden dimension을 순차적으로 128과 64로 설정하였다.

두 번째 encoder layer의 최종 hidden state ($h_N \in \mathbb{R}^{64}$)는 선형 투영 계층을 통해 32차원의 latent representation($z_N \in \mathbb{R}^{32}$)으로 압축된다. 이후 decoder는 latent representation으로부터 hidden dimension을 64와 128로 순차적으로 확장한 후, 최종 선형 계층을 통해 3개의 입력 변수(NIR, VIS, L/O)를 재구성하도록 설계하였다.

Table 5. Configurations of the baseline models (RNN-AE and LSTM-AE)

Parameter	RNN-AE	LSTM-AE
Input variables	3	3
Sequence length	1,080	1,080
Encoder_hidden	[128, 64]	[128, 64]
Encoder_layer	2	2
latent_dimension	32	32
Decoder_hidden	[64, 128]	[64, 128]
Decoder_layer	2	2
Output_dim	3	3
Params	~0.06M	~0.245M

Table 6. Training configuration shared by all models

Parameter	Value
Optimizer	Adam
Learning rate	1e-4
LR scheduler	Warmup
Batch size	32
Dropout	0.1
Max epochs	80
Early stopping patience	15

Table 6은 주요 학습 파라미터로 제안 모델과 비교 모델 모두 동일한 학습 데이터와 학습 설정에 대한 정의를 동일하게 유지하고 아키텍처만을 달리하였다. Fig. 9는 학습 과정에서의 Loss를 보여주며, Fig. 9(b) Multi-scale Only Model은 제안 모델(WTA)의 병렬 embedding 구조에서 CWT branch를 제거하고, 오직 Multi-scale convolution branch만을 사용한 구성으로, 다중 해상도를 갖는 1D convolution layer의 병렬 구조를 통해 token 표현을 구성한다. 이를 제안 모델과 비교함으로써, 시간-주파수 표현이 token 표현의 품질과 탐지 성능에 기

여 정도를 확인한다. 또한, Fig. 9(c)는 CWT Only 모델로 Multi-scale convolution branch를 제거하고 오직 Morlet wavelet transform을 embedding 차원 벡터를 만드는 방법을 활용한 제안 모델이다. 상세한 결과는 5.3절 Ablation study에서 기술한다.

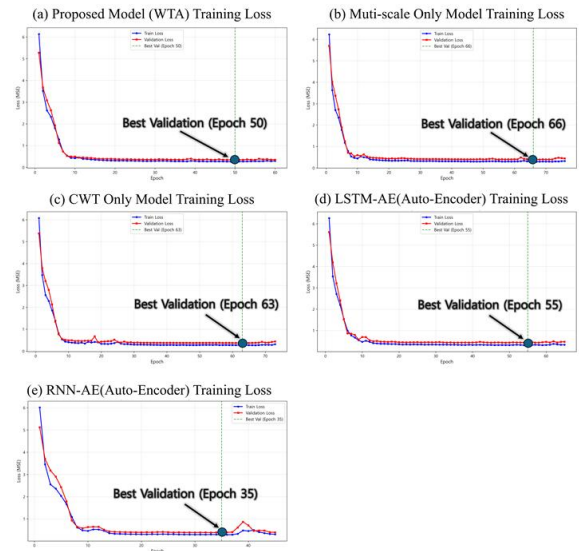


Fig. 9. Training and validation loss curves for (a) the proposed WTA model, (b) the multi-scale only model, (c) CWT only model, and (d) LSTM-AE, (e) RNN-AE

5.2 Evaluation Metrics

본 연구에서는 모델의 이상 탐지 성능을 정량적 평가를 위해 AUROC (Area under the ROC curve)와 AUPRC (Area Under the Precision-Recall curve) 그리고 Recall을 평가 지표로 활용하였다.

Recall은 훈련, 검증 데이터를 제외한 정상 데이터인 calibration set의 97th percentile로 산출된 threshold 값으로 계산하였다.

5.3 Analysis of experimental results

본 절에서는 제안 모델과 비교 모델 간의 이상 탐지 성능을 합성 이상 주입 실험을 통해 분석한다.

우선 3.3절에서 논의했던 Synthetic Anomaly Injection 결과를 바탕으로 Anomaly score 산출에 사용되는 Sliding window size(W)와 top- K 비율을 Grid search를 통해 도출되었다. Table. 1에 해당하는 9개의 시나리오와 Table. 2에 해당하는 5개의 이상 강도에 대한 탐색 격자는 $W \in \{100, 150, 200, 250, 300\}$ 이며, $K \in \{4, 6, 8, 10, 15, 20\}$ 설정하였다. 9개의 시나리오별로 5개의 이상 강도에 따른 AUROC, AUPRC 결과에 대한

평균값을 계산하였으며, 최적 지표로 해당 평균값이 제일 큰 파라미터를 최적 파라미터로 채택하였으며, Fig. 10은 그 결과를 보여준다.

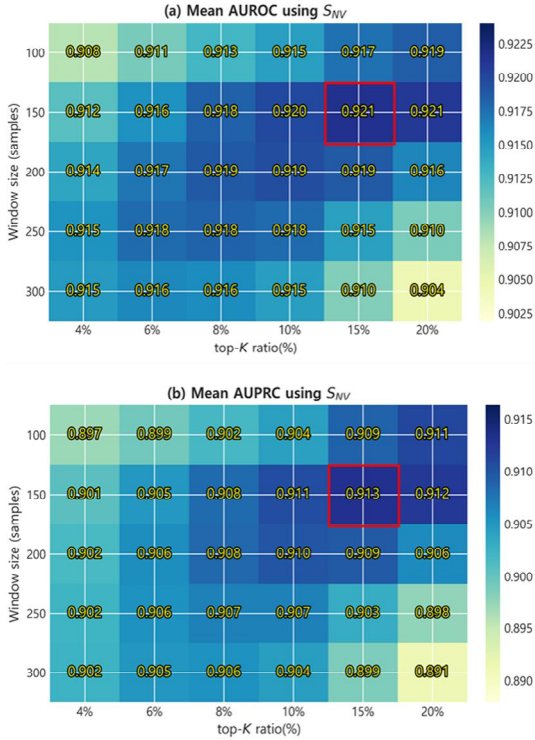


Fig. 10. Selecting the optimal sliding window size (W) and top- K %

가장 우수했던 지표로 sliding window size는 150, top- K 비율 15%로 선정되었으며, 이는 Anomaly score 계산 시 window 내부로부터 Point-wise error 중 상위 23개를 뽑아 평균값을 계산하게 된다. 이후의 모든 실험에서는 도출된 파라미터를 비교 모델과 ablation 구성에 동일하게 적용하여, Anomaly scoring 조건을 통제된 상태에서 모델에 의한 차이만을 비교하였다.

도출된 파라미터를 적용하여, 제안 모델(WTA)과 비교 모델 (RNN-AE, LSTM-AE)의 탐지 성능을 9가지 시나리오에 대해 비교하였다. Fig. 11은 각 시나리오에서 5가지 세기 레벨에 대해 산출한 AUROC와 AUPRC의 평균을 모델별로 보여준다.

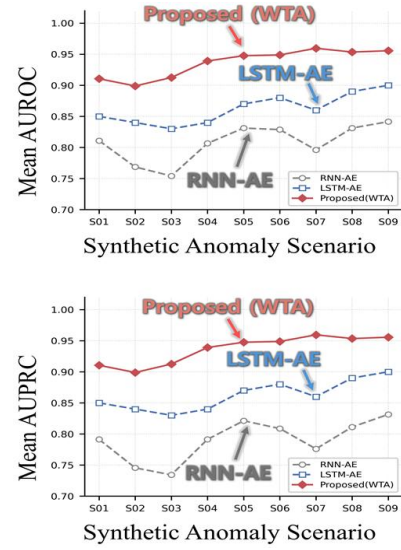


Fig. 11. AUROC and AUPRC of the proposed model(WTA) and RNN-AE, LSTM-AE for nine synthetic anomaly injection scenarios

전체 45가지 이상 주입 조건에 대한 평균 AUROC는 RNN-AE 0.8076, LSTM-AE 0.8622, 제안 모델 0.9362로 나타났으며, 제안 모델(WTA)이 두 비교 모델보다 높은 이상 탐지 성능을 보였다. 또한, 제안 모델(WTA)은 이상 주입 조건의 변화에도 전반적으로 안정적인 탐지 성능을 유지하였다.

시나리오별 경향을 살펴보면, 이상 주입 구간이 짧은 Short 시나리오(S01-S03)에서 모델 간 성능 차이가 상대적으로 크게 나타났다. RNN-AE와 LSTM-AE는 입력 시퀀스를 순차적으로 처리하고 이를 고정 길이의 latent representation으로 압축한 후 재구성하므로, 전체 시퀀스 중 짧은 구간에서 발생하는 국소적인 신호 변화가 잠재 표현에 충분히 보존되지 않을 수 있다.

반면 제안 모델은 multi-scale convolution을 통해 서로 다른 receptive field의 국소 시간 패턴을 추출하고, Morlet wavelet 기반 embedding을 통해 시점별 시간-주파수 특성을 함께 표현한다. 이러한 구조는 짧은 구간에서 발생하는 국소적 신호 변화에 대한 표현력을 높여 Short 시나리오에서도 상대적으로 높은 탐지 성능을 유지하는데 기여한 것으로 판단된다.

Fig. 12은 동일 포토다이오드 정상 신호에 특정 구간 이상 주입과 정상 신호에 대해 CWT branch를 통과했을 때의 결과이다. 정상 신호에서도 레이저 용접 시작과 끝지점은 과도구간이기 때문에 특정 주파수가 발생하지만, 안정 구간에서는 일부 국소적인 magnitude 변화가 나타난다. 하지만, 이상 주입 후에는 NIR 및 VIS 채널을 중심으로 정상 신호와 구별되는 magnitude 증가와 분포 변화가 뚜렷하게 관찰되며, 이러한 결과는 이상 구간에서의 신호 변

화가 시간-주파수 영역에서 국소적인 에너지 집중으로 나타나며, CWT branch가 이를 embedding 단계에서 포착함을 보여준다. 이는 앞서 Short 시나리오에서 제안 모델이 상대적으로 높은 탐지 성능을 보인 결과와 일관된다.

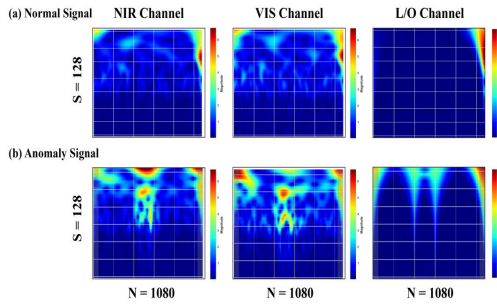


Fig. 12. CWT magnitude scalograms of the NIR, VIS, and L/O channels for (a) Normal signal and (b) the same signal after synthetic anomaly signal

5.4 Anomaly Detection on Laser welding Defects

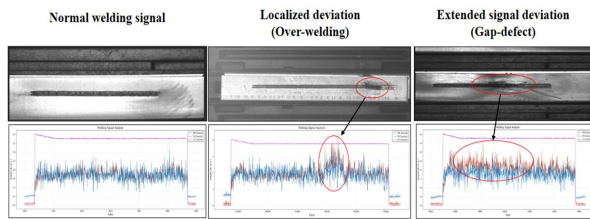


Fig. 13. Surface appearance and photodiode signals: (a) normal weld; (b) localized deviation (over-welding); (c) extended deviation (gap defect).

서론에서 기술한 바와 같이, 제안 모델은 특정 결함 유형의 신호 패턴을 학습하지 않고 정상 패턴으로부터의 이탈에 반응한다. 따라서 신호 변동의 발현 양상을 기준으로 양산 현장에서 발생하는 불량을 구분하였다.

선정된 국소 과용접은 누적 입열 현상에 의해 국소적 과잉 입열이 발생하며, 짧은 구간에 집중된 신호 증폭과 baseline 상승을 동반하는 특징을 갖는다. 갭 불량은 tab-to-tab 겹치기 용접에서 가압되지 고정 불량 등으로 모재 사이에 간극이 발생하면 계면에서의 열전달이 제한되어, 상대적으로 넓은 구간에 걸쳐 신호 변동의 심화를 동반하는 특징을 갖는다 [18].

제안된 모델의 현장 결함 탐지 성능을 검증하기 위해, 양산 현장 라인에서 취득한 두 가지 결함 유형인 각 불량 데이터는 국소 과용접 200개, 갭 불량 200개와 정상 신호 500개, 총 900개의 개별 용접 데이터를 사용하였다.

정상 데이터 500개는 이전 Calibration set에서 앞서 파라미터 조정 및 합성 이상 주입, Threshold 계산에 적용되지 않은 500개 데이터이며, 정상 calibration set의

97th percentile로 산출된 임계값(τ_{nv})을 적용한 결과, 오탐률은 0.022로 설계값 0.03 이하의 수준을 보였다.

Table 7. Detection performance on real welding defects evaluated by AUROC, AUPRC, and Recall

Model	Class	AUROC	AUPRC	Recall @ τ_{nv}
Proposed Model (WTA)	Localized Over-welding	0.92	0.92	0.90
	Gap-defect	0.94	0.93	0.91
LSTM-AE	Localized Over-welding	0.83	0.81	0.80
	Gap-defect	0.88	0.83	0.82
RNN-AE	Localized Over-welding	0.79	0.80	0.78
	Gap-defect	0.82	0.83	0.80

Table 4는 두 결함 유형에 대한 탐지 성능을 보여준다. 국소 과용접에 대해서는 AUROC 0.92, AUPRC 0.92를 기록하였으며, 갭 불량에 대해서는 AUROC 0.94, AUPRC 0.93으로 상대적으로 높은 값을 보였다. 임계값 τ_{nv} 에서의 Recall은 각각 0.90과 0.91를 보였다.

제안 모델은 두 결함 유형 모두에서 AUROC, AUPRC, Recall이 0.90 이상으로, 합성 이상 주입 조건뿐 아니라 실제 현장에서 발생하는 주요 불량에 대해서도 안정적인 탐지 성능을 보였다.

비교 모델과의 차이를 보면, 국소 과용접에서 제안 모델 (0.925)은 LSTM-AE(0.83) 및 RNN-AE(0.79) 대비 각각 0.10, 0.14 높은 AUROC를 기록하여, 짧은 구간에 집중된 이상에서 성능 차이가 두드러졌다.

5.5 Ablation study

본 절에서는 제안 모델의 hybrid embedding을 구성하는 두 branch의 기여를 확인하기 위해, 각 branch를 제거한 구성과 비교하였다. Multi-scale only는 CWT branch를 제거하고 시간 영역의 다중 해상도 표현만으로 embedding을 구성한 것이며, CWT only는 반대로 multi-scale convolution branch를 제거한 구성이다. 세 구성은 동일한 학습 데이터와 학습 설정에서 학습되었으며, anomaly scoring 파라미터($W=150, K=15\%$)와 평가 조건을 동일하게 유지하였다. 임계값은 각 구성의 정상 calibration set에서 97th percentile로 개별 산출하여, 동일한 오탐률 조건에서 Recall을 비교하였다.

Table 8은 실제 결함 데이터에 대한 ablation 결과를

보여준다. 두 branch 중 어느 하나를 제거한 경우 모든 지표에서 성능이 저하되었다. 국소 과용접에 대한 AUROC는 제안 모델 0.925에서 multi-scale only 0.874, CWT only 0.895로 감소하였으며, 갭 불량에 대해서는 0.941에서 각각 0.884, 0.900으로 감소하였다. Recall 또한 동일한 경향을 보였다.

Table 8. Ablation study on the embedding structure

Model	Class	AUROC	AUPRC	Recall @ τ_{nv}
Proposed Model (WTA)	Localized Over-welding	0.925	0.923	0.901
	Gap-defect	0.941	0.931	0.912
Multi-scale only	Localized Over-welding	0.874	0.863	0.854
	Gap-defect	0.884	0.865	0.835
CWT only	Localized Over-welding	0.895	0.881	0.862
	Gap-defect	0.900	0.880	0.854

단독 구성 간 비교에서는 CWT only가 multi-scale only 보다 두 결함 유형 모두에서 높은 성능을 보였다.

이는 레이저 용접 신호에서 이상에 의한 변화가 시간 영역의 진폭 변화뿐 아니라 주파수 구성의 변화로도 나타나며, 시간-주파수 표현이 이러한 변화를 포착하는 데 유효함을 시사한다. 다만 두 branch를 결합한 제안 모델이 두 단독 구성보다 모든 지표에서 높은 성능을 보였다는 점은, 시간 영역의 다중 해상도 국소 패턴과 시간-주파수 표현이 서로 중복되지 않는 상호 보완적 정보를 제공함을 의미한다.

VI. Conclusions

본 연구에서는 배터리 모듈 조립 공정의 tab-to-tab 레이저 용접을 이상 탐지를 위해 다변량 포토다이오드 신호 및 레이저 출력 신호를 활용한 비지도 기반 이상 탐지 모델을 제안하였다. 제안된 WTA 모델은 레이저 용접 광학 신호는 과도적 변동과 강한 비주기적 신호 특성을 보이기 때문에, 사전에 결함 유형과 특정 패턴을 사전 라벨링을 하기 어렵다는 문제와 기존의 시계열 패턴 학습 모델로는 국소적이고, 일시적으로 나타나는 이상 패턴을 효율적으로 식별하기 어려움을 갖고 있었다. 이에 본 연구에서는 기존 Transformer의 embedding 구조에 CWT branch와 Multi-scale dilated 1D convolution branch를 활용하여 레이저 용접 비정상적 신호에서 풍부한 특징 표현을 얻

고자 했다. 또한, 제조 현장에서 특정 구간에 불량 패턴에 대한 라벨링을 사전에 적용하기 힘들기 때문에 Synthetic Anomaly Injection에 대해 9가지의 시나리오를 적용하여 이상 주입 위치, 주입 기간을 변화와 5가지의 이상 강도를 변화시켜 제안된 WTA 모델의 성능 평가를 진행하였다.

이에 특히, 국소적인 이상 변화에서 기존 시계열 패턴 학습기반인 RNN-AE, LSTM-AE 보다 우수한 정확도를 보였다. 또한 실제 용접 현장의 주요 불량인 Localized Over-welding과 Gap-defect에 대해서는 제안된 모델이 결함에 대해 각각 AUROC 0.92과 0.94를 기록하며 이상 탐지에 우수한 탐지 성능으로 확인하였다.

본 연구에서 제안한 모델은 다변량 포토다이오드 센서 신호와 레이저 출력 신호를 활용하여, 복합적 물리 정보를 내포한 레이저 신호의 비정상성 및 비주기적 특성을 효과적으로 학습함으로써 기존 시계열 패턴 기반 모델 대비 우수한 성능을 보였다. 이를 실제 양산 공정의 주요 불량 유형인 국소 과용접 과 갭 불량에 대한 이상 탐지에 적용한 결과, 전기차 배터리 탭 간 용접의 전수 검사 단계에서 작업자의 품질 판정을 지원하는 실용적 모니터링 모델로 활용될 수 있음을 확인하였다.

향후 연구에서는 다음 세 가지 방향을 검토할 계획이다. 첫째, 본 연구는 self-attention의 이차 연산 부담을 완화하기 위해 시간축 down-sampling을 적용하였으나, 파라미터 효율성과 연산 비용을 함께 개선할 수 있는 대안적 구조를 검토할 필요가 있다. 둘째, 본 연구의 attention은 시간 축에 한정되어 있으므로, 변수 간 상관관계를 attention map으로 표현할 수 있는 variable attention을 병렬 구조로 결합하여 시간적 패턴과 변수 간 관계성을 동시에 학습하는 모델로의 확장을 검토한다. 셋째, 현재는 다채널 포토다이오드 신호만을 활용하고 있으나, CCD 또는 CMOS 카메라로 취득한 용융 풀 영상을 결합하는 멀티모달 접근을 통해 시계열 신호와 공간 정보를 함께 활용하는 방향을 목표로 한다.

ACKNOWLEDGEMENT

This work was supported through the Automotive Industry Technology Development Program (RS-2024-00507548, Development of the Informative Panel for Future Vehicles Through Adaptive Magnetic Patterns), funded by the Ministry of Trade, Industry and Energy (MOTIE) of the Republic of Korea and administered by the Korea Planning & Evaluation Institute of Industrial Technology (KEIT).

REFERENCES

- [1] I. Veza, M. Z. Asy'ari, M. Idris, V. Epin, I. M. R. Fattah, and M. Spraggon, "Electric vehicle (EV) and driving towards sustainability: Comparison between EV, HEV, PHEV, and ICE vehicles to achieve net zero emissions by 2050 from EV," *Alexandria Eng. J.*, Vol. 82, pp. 459-467, November 2023. DOI: 10.1016/j.aej.2023.10.020
- [2] S. Choi, S. Lim, A. Shan, J. Lee, T. G. Yun, and B. Hwang, "Tab-to-Busbar Interconnections in EV Battery Packs: An Introductory Review of Typical Welding Methods," *Micromachines*, Vol. 17, No. 1, pp. 2, January 2026. DOI: 10.3390/mi17010002
- [3] W. Cai, J. Wang, P. Jiang, L. Cao, G. Mi, and Q. Zhou, "Application of sensing techniques and artificial intelligence-based methods to laser welding real-time monitoring: A critical review of recent literature," *J. Manuf. Syst.*, Vol. 57, pp. 1-18, October 2020. DOI: 10.1016/j.jmsy.2020.07.021
- [4] F. Tenner, C. Brock, R. Hohenstein, Z. Zalevsky, and M. Schmidt, "Remote Optical Detection of the Fusion State in Laser Deep Penetration Welding," *Physics Procedia*, Vol. 41, pp. 515-519, 2013. DOI: 10.1016/j.phpro.2013.03.109
- [5] T. Sibillano, A. Ancona, V. Berardi, and P. M. Lugarà, "A Real-Time Spectroscopic Sensor for Monitoring Laser Welding Processes," *Sensors*, Vol. 9, No. 5, pp. 3376-3385, May 2009. DOI: 10.3390/s90503376
- [6] D. Y. You, X. D. Gao, and S. Katayama, "Review of laser welding monitoring," *Sci. Technol. Weld. Joining*, Vol. 19, No. 3, pp. 181-201, April 2014. DOI: 10.1179/1362171813Y.0000000180
- [7] S. A. Shevchik, T. Le-Quang, F. V. Farahani, N. Faivre, B. Meylan, S. Zanolì, and K. Wasmer, "Laser Welding Quality Monitoring via Graph Support Vector Machine With Data Adaptive Kernel," *IEEE Access*, Vol. 7, pp. 93108-93122, July 2019. DOI: 10.1109/ACCESS.2019.2927661
- [8] G. Liu, X. Gao, D. You, and N. Zhang, "Prediction of high power laser welding status based on PCA and SVM classification of multiple sensors," *J. Intell. Manuf.*, Vol. 30, No. 2, pp. 821-832, February 2019. DOI: 10.1007/s10845-016-1286-y
- [9] Z.-X. Ma, P.-X. Cheng, J. Ning, L.-J. Zhang, and S.-J. Na, "Innovations in Monitoring, Control and Design of Laser and Laser-Arc Hybrid Welding Processes," *Metals*, Vol. 11, No. 12, Article 1910, December 2021. DOI: 10.3390/met11121910
- [10] K. She, D. Li, K. Yang, M. Li, B. Wu, L. Yang, and Y. Huang, "Online Detection of Laser Welding Penetration Depth Based on Multi-Sensor Features," *Materials*, Vol. 17, No. 7, Article 1580, April 2024. DOI: 10.3390/ma17071580
- [11] Y. Niu, P. P. Gao, and X. Gao, "Recognition of DC01 Mild Steel Laser Welding Penetration Status Based on Photoelectric Signal and Neural Network," *Metals*, Vol. 13, No. 5, Article 871, May 2023. DOI: 10.3390/met13050871
- [12] Y. Zhang, D. You, X. Gao, N. Zhang, and P. P. Gao, "Welding defects detection based on deep learning with multiple optical sensors during disk laser welding of thick plates," *J. Manuf. Syst.*, Vol. 51, pp. 87-94, April 2019. DOI: 10.1016/j.jmsy.2019.02.004
- [13] E. C. Ozkat, "Photodiode Signal Patterns: Unsupervised Learning for Laser Weld Defect Analysis," *Processes*, Vol. 13, No. 1, Article 121, January 2025. DOI: 10.3390/pr13010121
- [14] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. Kaiser, and I. Polosukhin, "Attention Is All You Need," *arXiv preprint arXiv:1706.03762*, 2017.
- [15] G. Mattera and L. Nele, "Machine learning approaches for real-time process anomaly detection in wire arc additive manufacturing," *Int. J. Adv. Manuf. Technol.*, Vol. 137, pp. 2863-2888, 2025. DOI: 10.1007/s00170-025-15327-y
- [16] C. Torrence and G. P. Compo, "A Practical Guide to Wavelet Analysis," *Bull. Amer. Meteorol. Soc.*, Vol. 79, No. 1, pp. 61-78, January 1998. DOI: 10.1175/1520-0477(1998)079<0061:APGT WA>2.0.CO;2
- [17] P. Malhotra, A. Ramakrishnan, G. Anand, L. Vig, P. Agarwal, and G. Shroff, "LSTM-based Encoder-Decoder for Multi-sensor Anomaly Detection," *arXiv preprint arXiv:1607.00148*, 2016.
- [18] H. Kim, S. Kang, C. Kim, Y. H. Jang, and M. Kang, "AI-Driven Interfacial Gap Prediction in Overlapped Al/Cu Laser Weld Joint for Battery Applications," *IEEE Access*, Vol. 13, pp. 181302-181312, October 2025. DOI: 10.1109/ACCESS.2025.3621683

Authors



Jeong-Beom Hong received the B.S degree in Electronic Engineering from Gachon University, South Korea, in 2020, and the M.S. degree in Advanced Materials & Processing Engineering from Inha University,

South Korea, in 2022. He is currently a Ph.D. candidate in Advanced Materials & Processing Engineering at Inha University, Incheon, South Korea. His research interests include industrial AI, domain-specific AI, and materials simulation



Byung-Chul Ko received the Ph.D. degree in Metallurgical Engineering from Inha University, South Korea, in 1999. Dr. Ko joined the faculty of the Advanced Materials & Processing Research Center at Inha

University, Incheon, Korea, in 2024. His main research interests include industrial artificial intelligence, domain-specific artificial intelligence, materials simulation, and core industrial technology strategy.