

기술논문

다단계 딥러닝 기반 다이캐스팅 공정 불량 검출

이정수 * · 최영심 **

*가천대학교, **한국생산기술연구원

Fault Detection in Diecasting Process Based on Deep-Learning

Jeongsu Lee* and Youngsim Choi**

*Gachon University, Seongnam 13120, Republic of Korea

**Korea Institute of Industrial Technology, Siheung 11358, Republic of Korea

초 록

다이캐스팅 공정은 다양한 산업군의 인프라 역할을 수행하는 중요한 공정이지만, 높은 불량률로 인하여 관련 기업들의 수익성 및 생산성의 한계가 있는 상황이다. 이를 타개하기 위하여, 본 연구에서는 다이캐스팅 공정의 불량 검출을 위한 산업인공지능 기반 모듈을 구성하였다. 개발된 불량 검출 모듈은 제공되는 데이터의 특징에 따라서 3단계로 동작되는 모델로 구성된다. 1단계 모델은 비지도학습 기반 이상 검출을 진행하며, 레이블이 없는 데이터셋을 대상으로 작동한다. 2단계 모델은 반지도학습 기반으로 이상 검출을 진행하며, 양품 데이터의 레이블만 존재하는 데이터셋을 대상으로 작동하며, 3단계 모델은 소수의 불량 데이터가 제공된 상황의 지도학습 모델을 기반으로 작동한다. 개발된 모듈은 실제 다이캐스팅 양품 데이터를 바탕으로 96% 이상의 우수한 양품 검출 성능을 보였다.

핵심용어; 다이캐스팅, 정밀 주조, 불량 검출, 이상 감지, 산업인공지능

Abstract

The die-casting process is an important process for various industries, but there are limitations in the profitability and productivity of related companies due to the high defect rate. In order to overcome this, this study has developed die-casting fault detection modules based on industrial AI technologies. The developed module is constructed from three-stage models depending on the characteristics of the dataset. The first-stage model conducts fault detection based on supervised learning from the dataset without labels. The second-stage model realizes one-class classification based on semi-supervised learning, where the dataset only has production success labels. The third-stage model corresponds to fault detection based on supervised learning, where the dataset includes a small amount of production failure cases. The developed fault detection module exhibited outstanding performance with roughly 96% accuracy for actual process data.

Key words; Die-casting, Precision casting, Fault detection, Anomaly detection and Industrial AI.

- 이정수: 교수, 최영심: 수석연구원

Received: Sep. 30, 2022 ; Revised: Nov. 1, 2022 ; Accepted: Nov. 16, 2022

†Corresponding author: Youngsim Choi (KITECH)

Tel: +82-31-8084-8804, Fax: +82-31-8084-8603

E-mail: raycute@kitech.re.kr

Journal of Korea Foundry Society

2022. Vol. 42 No. 6, pp. 369~376

<http://dx.doi.org/10.7777/jkfs.2022.42.6.369>

pISSN 1598-706X / eISSN 2288-8381

© Korea Foundry Society, All rights reserved.

This is an Open-Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

1. 서론

스마트 팩토리 (Smart Factory) 및 ICT (Information & Communication Technology) 기술은 4차 산업 혁명의 주요한 기반 기술의 하나로서, 제조업 분야의 생산성 향상 및 경쟁력 제고를 유도하고 있다 [1,2]. 다이캐스팅 산업은 제조업 분야의 기반이 되는 뿌리 산업의 하나로서 첨단 제조 산업의 발전을 뒷받침하는 인프라의 역할을 하지만 스마트 팩토리를 비롯한 최신 생산 기술의 개발 및 적용은 미흡한 실정이다. 특히 다이캐스팅 공정의 노동집약적 특성과 낮은 영업이익률은 이러한 첨단 기술의 적용을 더욱 어렵게 만들며, 이러한 한계로 인하여 다이캐스팅 공정은 일반 제조 산업 대비해서 상대적으로 높은 5~10%가량의 불량률을 보이고 있다 [3,4].

본 연구는 다이캐스팅 산업의 고질적인 품질 불량과 제품 간 산포를 해결하기 위한 방안 중 하나로서 스마트팩토리 기술 중 하나인 산업인공지능 기반 불량 검출 기술 개발을 목적으로 한다. 최근 딥러닝 기술의 놀라운 발전은 다양한 분야에 걸쳐서 다양한 성공 사례를 만들어내고 있으며, 특히 생산 및 제조 분야에서는 장비의 이상 감지, 제품의 불량 검출 및 공정 설계 등 다양한 활용 분야에서 산업인공지능 기술의 발전을 만들어내고 있다 [5-7]. 그러나 관련된 연구의 폭발적인 증가에도 불구하고, 실제 생산 및 제조 현장에 관련 산업인공지능 기술이 적용된 사례는 드문 실정이다.

산업인공지능 기술 연구가 실제 생산 및 제조 현장에 적용되지 못한 주요한 이유 중 하나는 보통의 인공지능 기술은 다수의 명확하게 정돈된 데이터를 기반으로 작동하기 때문이다. 예를 들어 제품의 불량을 검출하고자 하는 경우에 데이터는 제품의 양품/불량 여부가 명확하게 기술된 레이블을 보유하고 있어야 한다. 그러나 산업 현장의 경우에는 관련된 물리적인 변수를 측정하기 어렵거나 어떤 물리 변수가 관련되어 있는지 알아차리기 어려운 경우가 많다. 또한 공정 데이터를 측정하기 위해서는 매우 높은 비용이나 노동력이 수반되는 경우가 대부분이다. 이러한 어려움을 해결하고 공정데이터 취득에 성공했다고 하더라도 최종적으로 데이터가 정돈되어 있지 않고 부정확한 경우가 대부분이다. 특히, 다이캐스팅 산업은 관련된 공정 데이터의 취득이 매우 어려우면서, 데이터 취득을 위한 기반이 구성되지 못한 경우가 대부분이다.

따라서 산업인공지능 기술을 다이캐스팅 공정에 적용하기 위해서는 상기한 데이터 취득의 어려움을 극복할 수 있는 방안이 필수적이다. 이를 위해서 본 연구는 제공된 공정데이터의 상태에 따라서 적합한 형태의 인공지능 알고리즘을 적용하여 불량을 검출하는 형태의 공정 진단 모듈을 개발하였다. 그리고 이러한 공정 진단 모듈을 실제 다이캐스팅 공정에 적용했을 때의 성능에 대한 분석을 진행하였다. 개발된 산업인공지능 기반 다이캐스팅 공정 불량 검출 모듈은 향후 스마트 다이캐스팅 팩토리 개발의 핵심 요소 기술로 활용될 것이라 기대된다.

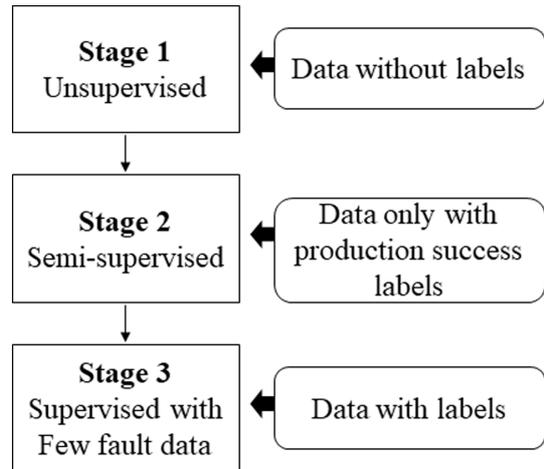


Fig. 1. Framework of the fault detection modules.

2. 산업인공지능 기반 불량 검출 모듈의 구성

2.1 불량 검출 모듈의 단계적 구성

본 연구의 공정 진단 모듈은 다이캐스팅 공정에서 취득되는 데이터의 특성을 다음의 3 단계로 구분한다 (Fig. 1). 가장 낮은 수준의 1단계 데이터는 제품의 양품/불량 상태에 대한 정보가 없는 상태에서 단순한 생산 데이터의 집합으로 제공된다. 1단계 불량검출 모듈이 성공적으로 개발될 경우 양산 현장에서 양품, 불량에 대한 구분이 필요 없이 데이터 취득 장비의 도입만으로 불량 검출이 가능하다. 이때의 데이터는 양품, 불량 레이블이 없이 단순한 숫자의 조합에 불과하며, 산업인공지능 알고리즘은 비지도 방식을 활용하여 데이터의 이상상태를 감지하는 역할을 해야 한다.

2단계는 보다 정돈된 형태의 데이터에 해당하며, 일부의 데이터에 대해서 확실한 양품인지 여부만 확인된 케이스이다. 보통의 다이캐스팅 제품 불량은 로트 (lot) 단위에서 제품 관리를 진행하며, 불량이 발생되었을 경우에 그 제품이 어떤 로트에 속해 있는지만 알 수 있다. 따라서 특정 로트에서 불량이 발생하지 않은 경우 그 로트에 속한 제품들에 대해서만 확실한 양품이라는 레이블을 부여할 수 있다. 이 경우 산업인공지능 알고리즘은 반지도 방식을 활용하여 제품의 이상을 검출하는 역할을 해야 한다.

3단계는 소수의 특정 제품에 대해서 확실한 불량이라고 말할 수 있는 경우이다. 보통의 공장에서는 다이캐스팅 제품을 모두 추적하지 않지만, 특수한 소수의 불량 제품에 대해서는 작업자가 특별히 그 이력을 추적하는 경우도 있다. 따라서 아주 소수의 불량 데이터가 제공될 수 있으며, 이 데이터를 기반으로 양품/불량 검출의 정확도를 높이는 것이 요구된다. 이 경우 산업인공지능 알고리즘은 소수의 불량 데이터를 적절하게 활용할 수 있는 형태로 개발되는 것이 바람직하다.

추가적으로 모든 제품에 대한 양품/불량 레이블이 제공되는

이상적인 경우를 생각해볼 수 있으나 실제 다이캐스팅 공정에서 모든 케이스에 대해서 양품/불량 레이블이 제공되는 경우는 현실적이지 않기 때문에 본 연구에서는 위 경우에 대응하는 알고리즘은 다루지 않고 있다.

2.2 1단계: 비지도 불량 검출 모듈

1단계의 비지도 불량 검출 모듈은 다이캐스팅 공정에 대해서 제품의 양산 시 금형의 온도 및 압력 취득 모듈만이 설치되어 있는 경우를 대응하는 모듈이다. 이 경우 추가적으로 생산된 제품의 이력 및 양품/불량 여부가 추가되지 않으며, 인공지능 모듈은 제공된 생산데이터만 보고 이상감지 알고리즘의 형태로 불량을 검출하게 된다. 이는 생산되는 대부분의 제품은 양품 (90~95%)이며, 소수의 불량 데이터 상의 이상치 (anomaly)로 검출될 것이라는 가정을 전제로 한다. 따라서 1단계 데이터를 대응하는 딥러닝 알고리즘은 제공되는 데이터를 비지도 (unsupervised) 방식을 기반으로 학습하여 제품의 이상유무를 판단하게 된다.

비지도학습 알고리즘은 기하학적 변환 (geometric transform)을 통한 자가 레이블 (self-label)을 바탕으로 한 이상치 검출을 기본으로 하며 [8], 딥러닝 알고리즘의 정확도를 향상시키기 위해 기하학적 변환에 대응될 수 있도록 시계열 데이터의 이미지 변환을 활용하였다. 시계열 데이터 이미지 변환에는 지문화 (fingerprint imaging)라고 명명하였으며, Fig. 2에 도시한 것과 같이 각 채널의 시계열 데이터를 극좌표계에서 각도와 길이로 변환하여 데이터의 채널에 해당하는 동심원의 형태로 구성하는 것을 의미한다. 이때 각 동심원의 순서를 변환하는 형태로 여러 종류의 지문화 변환을 활용하는 것이 가능하다.

다음으로 지문화 변환된 이미지 데이터는 기하학적 변환을 기반으로 자가 레이블을 형성하여 자가 학습을 진행하게 된다. 기하학적 변환이란 이미지 데이터를 기반으로 임의의 회전 변환을 진행하고, 딥러닝 네트워크는 적용된 회전 변환의 종류를 맞추는 학습을 진행하게 된다. 예를 들어 지문화 이미지를 바탕으로 90도, 180도, 270도, 360도 회전 변환을 진행하고 적용된 회전 변환의 종류를 맞추는 학습을 진행하는 것이다. 이러한 자가 학습을 통해서 딥러닝 네트워크는 기존에 학습했던 데이터와 유사한 형태의 데이터일 경우에 보다 높은 확신을 갖고 회전 변환의 종류를 예측하게 된다.

이상 감지의 관점에서 보면 보다 자주 접한 형태의 정상상태의 데이터의 경우는 예측의 확신도가 높고, 새롭게 등장하여 이상일 확률이 높은 형태의 데이터의 경우는 예측의 확신도가 떨어지게 된다. 딥러닝 네트워크의 확신도를 기준으로 이상 수치 (anomaly score)를 산출할 수 있으며, 1단계 알고리즘은 이를 기준으로 다이캐스팅 제품의 불량 검출을 진행한다.

2.3 2단계: 반지도 불량 검출 모듈

2단계의 반지도 불량 검출 모듈은 확실한 양품 레이블이 추가로 제공된 경우를 대응하는 모듈이다. 따라서 딥러닝 알고리즘은 단일 레이블 학습 (one-class classification)을 이용하여 반지도 (semi-supervised) 기반 불량 검출을 진행할 수 있도록 구성된다. 이를 위해서 딥러닝 기반 대표적인 one-class classification 알고리즘 중 하나인 Fence-GAN (FGAN) [9]이 활용되었다. 또한 이러한 딥러닝 기반 알고리즘이 이미지 형태의 데이터를 처리하는데 더욱 뛰어난 성능을 보인다는 점에 착안하여 시계열 데이터의 재귀적 특성을 강조하여

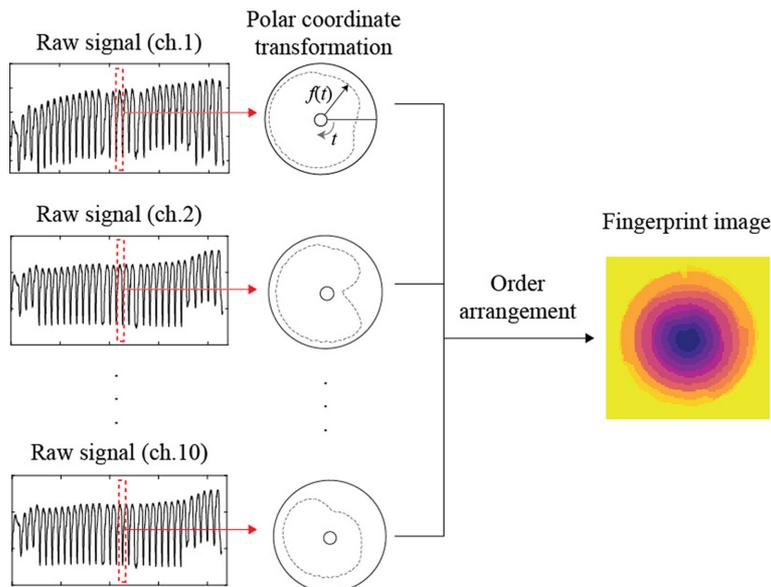


Fig. 2. Schematic diagram of the fingerprint imaging.

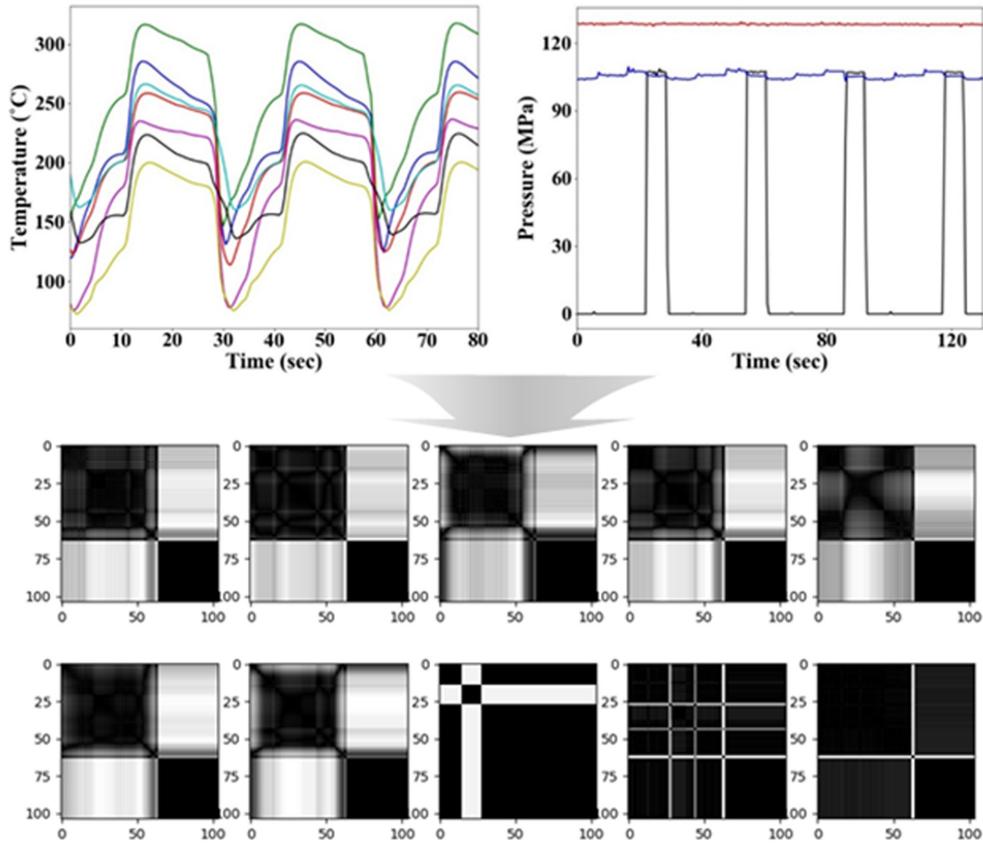


Fig. 3. Example of the recurrence plot of die-casting process data.

이미지를 형성하는 recurrence plot 및 markov transition field를 데이터 전처리 모듈로 활용하였다.

먼저 recurrence plot은 시계열 데이터의 재귀적 특성을 추출하여 공간 이미지 데이터를 형성하는 방식으로 다음의 과정을 통해서 얻을 수 있다 [10]. 시계열 데이터가 하기의 식 (1)의 형태를 지니는 시간 τ 간격으로 측정되는 m 개의 데이터라고 할 때 이미지 공간의 성분 $R_{i,j}$ 는 시계열 데이터의 i, j 번째 행의 차이를 통해서 다음의 식 (2)로 표현된다.

$$\vec{x}_i = (x_p, x_{i+p}, \dots, x_{i+(m-1)_i}) \quad (1)$$

$$R_{i,j} = \Theta(\varepsilon - \|\vec{x}_i - \vec{x}_j\|) \quad (2)$$

여기서 함수 Θ 는 단위 계단함수, ε 는 역치값을 의미하며 따라서 식 (2)의 의미는 i, j 번째 시계열 데이터의 차이가 역치 이상일 경우, 정의된 계단함수를 따라서 특정 값을 지니게 된다. Recurrence plot을 통한 다이캐스팅 데이터의 이미지 변환을 수행한 예시를 Fig. 3에서 확인할 수 있다.

다음으로 markov transition field는 시계열 데이터의 값을 기준으로 분위수 (quantile)를 상태 (state)로 지정한 후에 각 상태 간의 전이 확률 (markov transition field)을 행렬 형태

로 배치하여 이미지화하는 것이다. 이를 적용하기 위해서는 각 상태 x_t 를 분위수 기반 상태 q_t 로 변환한 후에 각 상태 간의 전이 매트릭스 $M_{i,j}$ 를 도출하는 과정을 거치게 된다. 따라서 전이 매트릭스는 시계열 데이터의 각 시점 i, j 간의 전이 확률 w_{ij} 로 정의되며 다음의 수식으로 표현된다.

$$M = \begin{bmatrix} w_{i,j|x_1eq_p, x_1eq_i} & \dots & w_{i,j|x_1eq_p, x_1eq_i} \\ w_{i,j|x_2eq_p, x_1eq_i} & \dots & w_{i,j|x_2eq_p, x_neq_i} \\ \dots & \dots & \dots \\ w_{i,j|x_neq_p, x_1eq_i} & \dots & w_{i,j|x_neq_p, x_neq_i} \end{bmatrix} \quad (3)$$

$n \times n$ 매트릭스 $M_{i,j}$ 는 동일한 크기의 이미지로 변환될 수 있다. Markov transition field는 분위수 기반으로 지정된 상태의 전환 확률을 통해서 표현되므로 실제로 시계열 데이터 상의 각 시점에서의 시간에 따른 변화정보가 소실되나, 각 상태들의 전환확률을 통하여 표현되므로 미세한 특징 변화를 민감하게 표현할 수 있는 장점이 있어, 시계열 데이터의 이미지 변환으로 활용 가능하다.

마지막으로 recurrence plot 및 markov transition field를 통해 변환된 이미지는 FGAN을 활용하여 양품/불량 판정을 진행하게 된다. FGAN은 생성모델 (generative model)의 대표적인 알고리즘인 GAN (generative adversarial networks)

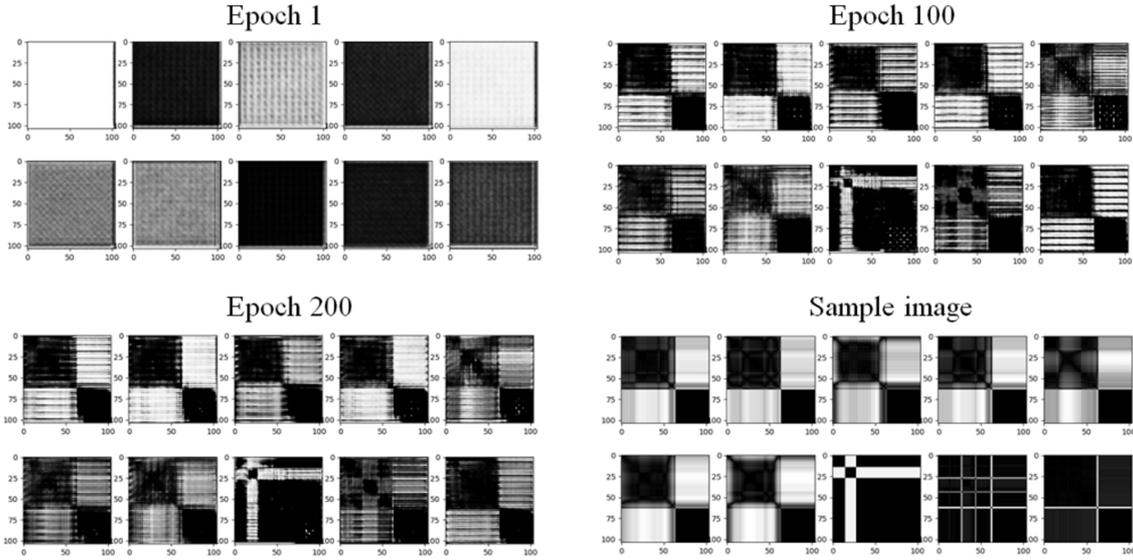


Fig. 4. Images generated by Fence-GAN by training epochs.

을 활용한 one-class classification 알고리즘이다. 여기서 GAN 구조는 임의의 랜덤 변수 Z 를 입력으로 가짜 데이터를 합성하는 생성자 (generator, G_θ)와 가짜인지 진짜인지 구별하는 구분자 (discriminator, D_ϕ)의 적대적 학습으로 구성된다. FGAN은 제품의 양품/불량 판정에 GAN을 이용하기 위하여 가짜 샘플들의 특성을 나타내는 매니폴드가 실제 제품의 경계에 위치하도록 학습한다. 이를 위해서 FGAN은 generator가 실제 데이터의 매니폴드 주변에 가짜 데이터를 형성하도록 encirclement loss (EL)와 dispersion loss (DL)를 사용하며 두 Loss의 상대적인 크기는 regularization parameter β 를 통하여 조정된다. (식 (4)) 최종적으로 이때의 매니폴드 경계선을 기준으로 양품/불량을 판정하게 되며, 이를 위하여 식 (5)로 표현되는 discriminator score를 anomaly score로 바로 활용한다. 생성 모델은 Input data를 기반으로 이를 모방하는 데이터를 생성하므로, 학습 과정에 따라서 적절하게 데이터의 생성이 이루어졌는지 확인이 우선적으로 필요하며 이를 Fig. 4에 도시하였다. 따라서 Fig. 4에서 FGAN의 학습 과정에 따라 생성모델로부터 생성된 recurrence plot의 샘플 이미지를 확인할 수 있다. 여기서 N , γ 는 각각 이용한 데이터의 개수 및 Generator와 Discriminator 간의 Loss regularization parameter를 의미한다.

$$LFGAN_{\geq \text{rator}}(G_\theta, D_\phi, Z) = EL \times \beta + DL \quad (4)$$

$$LFGAN_{\text{discriminator}}(G_\theta, D_\phi, \gamma, Z) = \frac{1}{N} \sum_{i=1}^n [-\log(D_\phi(X_i)) - \gamma \log(1 - D_\phi(G_\theta(z_i)))] \quad (5)$$

2.4 3단계: 소수 불량 데이터 활용 모듈

3단계의 불량 검출 모듈은 소수의 불량 데이터를 활용하여 기존의 불량 검출 알고리즘의 양품/불량 판정 성공률을 향상시키는 것을 목표로 한다. 일반적으로 제조업의 불량 데이터는 그 획득이 매우 어려워 소수의 데이터 수집만이 가능하지만, 불량 검출을 위한 중요한 정보를 제공한다. 그러나 일반적인 딥러닝 알고리즘은 고차원으로 주어진 다수의 데이터를 활용하여 정확한 판단을 내리는 것에 특화되어 있으며, 소수의 데이터 기반으로는 작동하기 어렵다.

따라서 소수의 데이터를 기반으로 정보를 획득하기 위해서는 기존에 지니고 있는 도메인 지식을 이용하여 활용할 데이터의 차원을 감소시키는 것이 필요하다. 이를 위해서는 다이캐스팅 공정의 불량에 대한 이해가 필요하다. 다이캐스팅 제품에는 다양한 유형의 불량이 존재하는데 대표적으로는 표면 결함, 수축기포 결함, 크랙, 내부 기포 결함을 예로 들 수 있다 [11]. 이러한 제품 불량의 원인을 크게 구분해보면 용탕 (molten metal)의 온도 또는 흐름 불량으로 구분할 수 있고, 물리적으로 용탕의 온도 및 압력을 통해서 각 불량의 발생을 예측할 수 있다.

먼저 온도 측면에서의 불량의 발생 상황을 추측해보면 실제 공장에서 제품이 생산될 경우 수백개의 제품이 연속적으로 생산이 되나, 작업 조건, 기기 부하, 작업자 상태 등의 조건에 따라서 각 생산 간 대기 시간의 편차 또는 일시적인 기기 가동의 중단이 발생한다. 이를 금형 온도의 측면에서 생각하면 생산 중단 시 금형의 온도가 정상조건에 비해서 급락했다가, 제품의 생산이 재개되면서 정상 온도를 회복하는 것을 Fig. 5를 통해서 확인할 수 있다. 또한 반복적인 생산으로 인하여 금형의 온도가 특정 정상상태 조건에 가까워졌

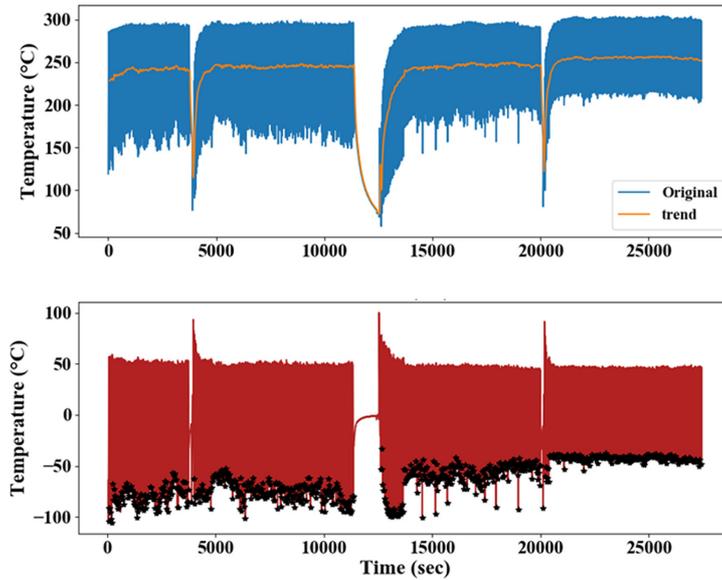


Fig. 5. Mold temperature over process cycles.

을 경우에는 제품의 성형 조건 및 냉각 조건에 따라서 랜덤 편차를 보이게 된다.

따라서 금형 온도의 관점에서 생산 편차는 두 가지로 요약할 수 있다. 첫 번째는 반복 생산의 열량 누적으로 표현되는 평균 온도 관점의 산포와 두 번째는 개별 생산 시 랜덤하게 발생하는 공정의 조건 산포로 구분이 가능하다. 이 점에서 착안하여 제공된 공정데이터의 차원을 축소하기 위하여 다음의 전처리를 수행한다. 먼저 FFT (Fast Fourier Transform) 분석을 통하여 제품 생산 주기를 추출하고 이동 평균 (moving average)을 수행하여 금형의 추세선 (trend line)을 통한 생산 별 평균 온도를 얻을 수 있다. 다음으로는 이 추세선과 온도 데이터의 차이를 이용하여 최대, 최소 온도 편차를 얻을 수 있으며 이 데이터를 인공지능 알고리즘에 투입하는 입력 데이터로 활용한다.

진술한 금형 온도 분포의 전처리 과정을 거쳐서 측정된 공정데이터들을 금형 추세 온도 효과와 제작 사이클 효과로 분리하였다. 다음으로 각 제품의 생산 공정 당시의 온도 데이터를 대상으로 공정 시간, 추세 평균 온도, 사이클 평균 온도, 사이클 최대 온도, 사이클 최소 온도의 다섯 개 변수를 7개의 온도 측정 채널을 대상으로 추출하여 총 35개의 입력 변수를 선정할 수 있다. 압력 관점에서는 우선적으로 두 개의 어큐뮬레이터에 인가되는 압력 채널을 모니터링하여 각 공정 당 평균, 최대, 최소 압력의 총 6개의 데이터를 추출하였다. 또한 압력 강화 (pressure intensification) 과정을 모니터링하여 압력 인가 시간, 압력 증가/감소율, 평균 및 최대 압력의 5개의 변수를 추가하였다. 마지막으로 제품의 생산 시간을 추가하여 총 47개의 변수를 활용하여 저차원화된 데이터를 활용하여 제품의 양품/불량을 판정하는 딥러닝 모델을

Table 1. Process variables adopted in 3rd phase fault detection module

No.	Data Source	Variables
1	Time	Cycle time
2		Pressure applied time
3		Pressure increase rate
4	Hydraulic injection pressure	Pressure decrease rate
5		Cycle-averaged pressure
6		Maximum intensification pressure
7		Cycle-averaged pressure
8	Accumulator pressure of the intensification	Maximum pressure
9		Minimum pressure
10		Cycle-averaged pressure
11	Accumulator pressure of the fast filling	Maximum pressure
12		Minimum pressure
13		Elapsed time to the highest temperature
14		Cycle-averaged temperature
15	Temperature #1	Mean level of fluctuation
16		Maximum fluctuation
17		Minimum fluctuation
		...
47	Temperature #7	Minimum fluctuation

개발하였다. 도메인 지식을 활용하여 저차원화된 데이터의 목록은 Table 1을 통해서 확인할 수 있다.

3. 다이캐스팅 데이터 활용 결과

3.1 다이캐스팅 공정 데이터 수집

다이캐스팅 공정 데이터는 한국생산기술연구원에서 자체 개발한 다이캐스팅 스마트 팩토리 플랫폼 기술을 [12] 활용하여 측정하였다. 다이캐스팅 제품의 불량은 생산이 완료된 후 후가공 과정에서 대부분 검출이 되며, 이를 위해서는 제품 이력의 추적 관리가 핵심이 된다. 그러나 대부분의 다이캐스팅 제품은 이력의 추적 관리가 불가능하기 때문에 다이캐스팅 금형의 밀핀을 이용하여 대량 생산되는 다이캐스팅 주조품에 대하여 생산 일련번호가 추적되는 시스템을 적용하였다. 관련된 시스템은 실제 제품의 생산되는 다이캐스팅 공정에 적용하여 총 1500여 건의 양산 공정 데이터를 수집하였으며, 그 중 불량은 70건이 발생하였다. 따라서 총 1430건의 양품과 70건의 불량 중 랜덤으로 양품 50건, 불량 50건을 추출하여 100건의 검증용 데이터셋을 구성하고, 나머지 데이터는 학습에 이용하였다. 위 3단계로 개발된 공정 불량 활용 모듈을 적용하여 각 모듈의 정확도를 검증하여 보고하였다.

전술한 바와 같이 1단계 모듈에는 양품/불량에 대한 정보 없이 비지도 기반으로 데이터를 투입하였으며, 2단계 모듈에는 양품 레이블만 제공된 반지도 모듈, 그리고 3단계 모듈을 대상으로 소수의 불량 레이블이 제공된 데이터셋을 활용하였다. 불량 검출 모듈의 성능 비교를 위해서는 AUROC (Area under receiver operating curve) 지표를 활용하였다. AUROC는 True Positive Rate (TPR, 전체 양품 중 양품으로 정확히 판단한 비율)와 False Positive Rate (FPR, 전체 불량 중 양품으로 잘못 판단한 비율)을 각각 y축, x축으로 하는 그래프의 아래 면적을 의미한다. 따라서 불량 검출 모듈이 이상적으로 양품, 불량을 모두 정확히 판단했을 경우 면적 1을 갖게 되고, 아래 면적은 %의 개념으로 표현될 수 있다.

3.2 불량 검출 모듈 결과 요약

전술한 시계열 데이터 이미지 변환 및 self-labeled training 기반 불량 검출 모듈을 이용하여 다이캐스팅 제조 데이터에 대한 불량 검출 테스트를 진행하였다. 다이캐스팅 공정데이터는 총 10개의 채널에서 계측한 온도 및 압력 데이터로 구성하며, 10개의 채널 순서를 바꾸어 가며 이미지 변환 및 label을 진행하여 채널 순서 변환을 학습하였다. 그 결과 본 연구에서 제안하는 비지도 불량 검출 모델은 이미지 변환 종류 개수에 따라서 정확도가 변화하는 것을 확인하였고, 이미지 변환 종류의 개수를 포함한 하이퍼 파라미터 (hyperparameters)를 최적화하였다. 최종적으로 30종의 변환을 이용하였을 때 AUROC 기준 0.937로 가장 높은 성능을 보였다. 이는 양품/불량이 제공되지 않은 데이터를 기반으로 한 결과로, 제품의 양품, 불량 판단을 진행하지 않고 단지 데이터 측정 시스템만을 다이캐스팅 공정에 설치하여 대략 93%의 불량을 검출할 수 있다는 것을 의미한다.

Table 2. Performance of fault detection modules in each stage

	Algorithm	AUROC
Stage 1	Unsupervised (Geometric transformation)	93.70%
Stage 2-1	Semi-supervised (FGAN with RP)	91.76%
Stage 2-1	Semi-supervised (FGAN with MTF)	94.04%
Stage 3	Supervised (Fully connected neural networks)	96.9% (Accuracy)

다음 단계로 특정 제품군에 대해서만 확실한 양품이라는 조건이 주어진 Stage 2에 대해서 불량 검출 모듈을 적용하였다. 따라서 양품 50, 불량 50으로 구성된 테스트 데이터를 제외하고 남은 양품 데이터를 바탕으로 반지도 학습 모듈을 학습시켰다. 반지도 학습 모듈의 전처리 모듈로서 시계열 데이터를 전술한 RP (Recurrence plot)와 MTF (Markov transition field)를 배경으로 처리한 후 Geometric transformation 기반 불량 검출 모듈에 투입하였다. 그 결과 RP를 활용할 경우는 91.76%, MTF를 활용하였을 때는 94.04%로 Stage1 모듈에 비해서 미세하게 증진된 결과를 얻을 수 있었다 (Table 2).

마지막 3단계 불량 검출 모듈은 소수의 불량 데이터를 활용하는 모듈로서, 지도학습 형태의 인공지능 모듈을 통해서 불량을 검출하는 모듈이다. 인공지능 모듈로서는 소수의 저차원 데이터를 처리하기에 적합한 단순한 FCNN (Fully Connected Neural Networks)을 활용하였다. FCNN은 2개의 hidden layer를 활용하였으며, 각 hidden layer는 200개의 뉴런을 지니도록 구성되었으며 마지막 layer에서는 양품/불량에 대한 판단을 내리는 layer에 해당한다. 3단계 불량 검출 모듈의 경우 소수의 불량 레이블 데이터에 대한 학습이 필요하다. 그러나 전체 불량 데이터는 70개 밖에 되지 않기 때문에 이 경우 랜덤 샘플링을 통해서 테스트 데이터 80개, 불량 데이터 20개로 구성된 총 100개의 테스트 데이터를 선정하고 나머지 1400개의 데이터를 통하여 학습을 진행하였다. 따라서 불량 검출 모듈은 총 50개의 불량 케이스에 대한 학습을 진행하게 되었다. 그리고 이 과정을 10번 반복하여 총 800개의 양품 데이터, 200개의 불량 데이터에 대한 결과를 진단하였다. 그 결과 Table 2에 나타난 것처럼 불량 검출 모듈은 정확도 (Accuracy) 기준으로 96.9%의 불량 검출 성공률을 보였다.

4. 결 론

본 연구에서는 대표적인 정밀주조 공법 중 하나인 다이캐스팅 공정을 대상으로 제품의 품질 불량을 검출하는 알고리즘을 개발하였다. 개발된 알고리즘은 독립적으로 3단계로 동작하도록 설계되어 각 알고리즘이 단계별 적용에 영향을 주

지 않고, 양산 현장의 데이터 취득 특성에 따라서 적용 변화할 수 있도록 설계되었다. 1단계는 제품의 양품, 불량에 대한 정보가 없이 단순 데이터만 취득하여 비지도 이상감지 방식으로 불량을 검출하는 Geometric transformation 기반 방식, 2단계는 제품의 양품 레이블만 제공되어 반지도 one-class classification 방식으로 작동하도록 설계된 알고리즘으로, 시계열 데이터 이미지화를 위한 recurrence plot 및 markov transition field 및 fence-GAN 알고리즘을 적용하였다. 이때는 다이캐스팅 공정에서 불량이 발견되지 않은 로트 (lot)가 존재하여, 완벽하게 양품인 데이터만 선별 가능한 상황을 목적으로 설계되었다. 마지막 3단계는 수십개 수준의 아주 적은 숫자의 불량 데이터가 레이블과 함께 적용되어 이용 가능한 상황을 산정하고 설계되었다. 본 연구에서 제시한 알고리즘은 각각의 특화된 상황에서 우수한 불량 검출 성능을 보였으며, 향후 다이캐스팅 제조 환경의 스마트화 및 불량 저감에 기여할 것으로 기대된다. 또한 향후에는 불량의 종류에 대한 추가적인 구분과, 불량에 대한 정보를 활용하여 주조 환경을 최적화하는 연구가 이루어질 수 있을 것이라 생각된다.

감사의 글

본 연구는 한국생산기술연구원 제조혁신지원사업의 지원을 받아 수행하였습니다.

References

- [1] Wang S, Wan J, Li D and Zhang C., *Int. J. Distributed Sensor Networks*, **12**(1) (2016) 3159805.
- [2] Lucke D, Constantinescu C and Westkämper E., *Manufacturing systems and technologies for the new frontier*, Springer, London (2008) 115.
- [3] Shanmugaraja M, Nataraj M and Gunasekaran N., *Int. J. Quality and Innovation*, **1**(3) (2011) 259.
- [4] T.W. Kim, C.K. Jin, I.K. Jeong, S.S. Lim, J.C. Mun, C.G. Kang, H.Y. Seo and J.D. Kim, *Metals*, **5**(2) (2015) 706.
- [5] Wang J, Fu P and Gao RX, *J. Manufacturing Systems*, **51** (2019) 52.
- [6] Lin H, Li B, Wang X, Shu Y and Niu S, *J. Intelligent Manufacturing*, **30**(6) (2019) 2525.
- [7] M.S. Kim, M.Y. Lee, M.J. An and H.C. Lee, *J. Intelligent Manufacturing*, **31**(5) (2020) 1165.
- [8] Golan I and El-Yaniv R., *Adv. neural information processing systems*, (2018) 31.
- [9] Ngo PC, Winarto AA, Kou CKL, Park S, Akram F and Lee HK, *IEEE 31st Int. Conference on tools with artificial intelligence* (2019) 114.
- [10] March TK, Chapman SC and Dendy RO, *Physica D: Nonlinear Phenomena*, **200**(1-2) (2005) 171.
- [11] Butler, B., "Die-casting handbook", USA, NADCA (2001).
- [12] J. Lee, Y.C. Lee and J.T. Kim, *J. Materials Processing Technology*, **290** (2021) 116972.