기계 학습을 활용한 마그네슘 합금의 통전 비선형 온도 예측

유진영 · 이명재 · 문영훈 · 노유정 · 이태경*

부산대학교 기계공학부

Prediction of Electropulse-Induced Nonlinear Temperature Variation of Mg Alloy Based on Machine Learning

Jinyeong Yu, Myoungjae Lee, Young Hoon Moon, Yoojeong Noh, and Taekyung Lee*

School of Mechanical Engineering, Pusan National University, Busan 46241, Republic of Korea

Abstract: Electropulse-induced heating has attracted attention due to its high energy efficiency. However, the process gives rise to a nonlinear temperature variation, which is difficult to predict using a traditional physics model. As an alternative, this study employed machine-learning technology to predict such temperature variation for the first time. Mg alloy was exposed to a single electropulse with a variety of pulse magnitudes and durations for this purpose. Nine machine-learning models were established from algorithms from artificial neural network (ANN), deep neural network (DNN), and extreme gradient boosting (XGBoost). The ANN models showed an insufficient predicting capability with respect to the region of peak temperature, where temperature varied most significantly. The DNN models were built by increasing model complexity, enhancing architectures, and tuning hyperparameters. They exhibited a remarkable improvement in predicting capability at the heating-cooling boundary as well as overall estimation. As a result, the DNN-2 model in this group showed the best prediction of nonlinear temperature variation among the machine-learning models built in this study. The XGBoost model exhibited poor predicting performance when default hyperparameters were applied. However, hyperparameter tuning of learning rates and maximum depths resulted in a decent predicting capability with this algorithm. Furthermore, XGBoost models exhibited an extreme reduction in learning time compared with the ANN and DNN models. This advantage is expected to be useful for predicting more complicated cases including various materials, multi-step electropulses, and electrically-assisted forming.

(Received April 2, 2020; Accepted April 28, 2020)

Keywords: machine learning; magnesium; electropulse; electrically-assisted forming

1. 서 론

통상적인 마그네슘 합금은 슬립계의 제한에 따른 낮은 상온 성형성 때문에 냉간 성형이 어렵다 [1]. 따라서 상기 소재의 소성 가공은 비가치면 슬립이 활성화되는 180 °C 이상의 온도에서 온간 혹은 열간 성형을 통해 수행된다 [2]. 이를 위해 해당 소재를 전기로 등 고온 환경에서 일정 시간 가열 후, 상온 환경에서 소재가 냉각되기 전에 압연

이나 단조 등의 소성 가공을 적용하는 것이 가장 일반적인 성형 공정이라 할 수 있다. 그러나 이러한 방식은 소재 가열을 위한 고온 환경을 별도로 구축해야 하므로 그에 따른 비용 증가 및 에너지 효율 저하가 불가피하다 [3].

이러한 문제를 해결하기 위한 방안은 크게 두 방향으로 나뉜다. 첫 번째는 미세조직이나 합금 성분을 제어하여 상온 성형성이 증진된 신소재를 개발하는 것이고, 두 번째는 온간 혹은 열간 성형 공정 자체의 효율을 높이는 것이다. 후자와 관련하여 2000년대에 들어 새롭게 주목받고 있는 가공법이 통전 성형(electrically-assisted forming)이다 [4]. 통전 성형은 금속 소재에 전류를 인가할 때 발생하는 줄열(Joule heating)을 활용한다. 따라서 기존 가열 방식 대비

- 유진영: 석, 박사 통합과정, 이명재: 석사과정, 문영훈: 교수, 노유정: 부교수, 이태경: 조교수

*Corresponding Author: Taekyung Lee

[Tel: +82-51-510-2985, E-mail: taeklee@pnu.ac.kr]

Copyright © The Korean Institute of Metals and Materials

에너지 효율이 높고, 금속 가열이 가능할 뿐 아니라, 고온 환경을 구축할 필요가 없어 공정 단축의 강점도 가진다 [5,6]. 초기의 통전 성형은 금속 소재에 전류를 연속적으로 인가하는 방식을 사용하였으나, 소재에 따라 오히려 연성을 감소시키는 사례가 보고되었다 [7,8]. 이러한 문제를 해결하기 위한 대안으로 최근의 통전 관련 연구는 전류 펄스를 주기적으로 인가하는 방식을 사용하고 있다 [9,10].

통전 성형에 있어 전류 펄스는 전류 밀도로 표현되는 크기(magnitude), 단일 펄스의 유지 시간을 나타내는 길이(duration) 및 펄스 사이의 시간 간격인 주기(period)로 정의할 수 있다 [11]. 상기 세 가지 특성은 전류 펄스의 형태를 결정할 뿐 아니라, 전류 펄스가 인가된 소재의 미세조직 및 변형 특성에 직접적인 영향을 끼친다. 예를 들어 Salandro 등[11]에 따르면 알루미늄 합금에 1초 길이의 90 A/mm² 전류 펄스를 분당 1회 인가 시 연성이 54.3 % 증진되는 반면, 60 A/mm² 펄스를 분당 2회 인가할 경우 연성은 겨우 2.2 % 향상되는 것에 그친다. Lee 등[12]은 전류 펄스에 의한 마그네슘 합금의 스프링백 감소 효과를 연구하였는데, 전류 펄스 길이를 두 배 증가할 때보다 전류 펄스 크기를 두 배 증가할 때 더 큰 감소 효과를 확인하였다. 이렇듯 소재의 특성을 개선하기 위해서는 전류 펄스의 형태를 최적화하는 작업이 불가피하다. 통상적인 최적화 방안은 상기 세 변수로 가능한 조합을 모두 시험해보는 격자 탐색(grid search)이다. 그러나 통전 가열은 그 특성상 소재의 종류, 통전 가열 이전 가공 이력, 초기 미세조직, 초기 물성 등 추가로 고려해야 할 변수가 많은 탓에, 격자 탐색은 매우 제한된 범위 내에서만 유효하다고 할 수 있다.

본 연구 그룹은 이러한 문제를 해결하기 위한 새로운 접근법으로, 기계 학습(machine learning)을 활용하고자 한다. 기계 학습은 특유의 빠른 예측 속도 및 정확성으로 인해 최근 재료 분야 전반에 활발히 활용되고 있다 [13-15]. 이러한 견지에서 Agrawal과 Choudhary[16]는 빅데이터 기반 기계 학습을 재료 과학의 제4차 패러다임으로 정의하기도 하였다. 본 연구는 기계 학습 기반 통전 성형 공정을 다루는 최초의 연구로, 우선 단일 펄스에 대한 크기 및 길이를 변수로 하여 수행되었다. 두 종류의 기계 학습 아키텍처로부터 9가지의 모델을 설계하여 실험 결과를 학습시킨 후, 전류 펄스에 따른 비선형 온도 변화에 대한 예측 정확성 및 학습 시간을 비교하였다.

2. 실험 방법 및 데이터 전처리

본 연구에서 사용한 소재는 상용 마그네슘 합금인

ZK60A-T5 압출재이다. 주요 성분의 화학 조성은 중량 비로 Mg-5.29Zn-0.51Zr-0.10Mn이다. 통전 시험을 위해 이 압출재로부터 길이 70 mm, 폭 10 mm, 두께 2 mm의 판상 시편을 채취하였다. 이 판상 시편의 길이는 압출 방향과 평행하다. 통전 시험 전 모든 시편의 표면을 #800 및 #1200 사포로 연마하여 산화막을 제거하고, 열화상 스프레이를 도포하여 방사율을 보정하였다. 통전 시험은 교류 전원, 전류 펄스 제어부, 절연 처리된 시편 고정체로 구성된 자체 제작 장비로 수행되었다. 해당 장비를 통해 전류 밀도 42.4-73.1 A/mm², 통전 시간 1,200-3,700 ms 범위의 전류 펄스를 1회 인가하여 각 시편을 순간적으로 가열 후 공랭하였다. 이때 시편에 인가된 전류는 클램프 미터를 통해 측정하였으며 통전에 의한 온도 변화는 적외선 열화상 카메라로 녹화 후 분석하였다.

12회의 통전 시험 결과 11,191개의 인스턴스가 확보되었다. 데이터 단위의 불일치에 따른 모델 성능 저하를 방지하기 위해, 기계 학습 전 각 특성의 평균(μ) 및 표준편차(σ)를 활용하여 모든 인스턴스를 다음과 같이 표준화(standardization) 처리하였다:

$$z = (x - \mu) / \sigma \quad (1)$$

기계 학습에 사용되는 통상적인 접근법은 주어진 샘플을 무작위로 세 그룹으로 나누어 각각 훈련 세트, 평가 세트, 테스트 세트로 활용하는 것이다. 그러나 본 연구의 경우 특정 펄스 크기 및 길이를 설정했을 때 그에 따른 온도 곡선을 온전히 얻는 것이 목적이므로 다소 다른 접근법이 사용되었다. 즉, 12회의 실험 데이터 중 전류 밀도 59.5 A/mm², 통전 시간 1,700 ms에 해당하는 모든 인스턴스(672개)를 테스트 세트로 지정하였다. 이 테스트 세트는 모든 학습 과정에서 완전히 배제되고 오직 최종 평가 과정에서만 사용되었다. 남은 10,519개의 인스턴스를 섞은 후 8:2 비율로 나누어 각각 훈련 및 평가 세트로 지정하였다. 이때 실험 재현성을 확보하기 위해 고정 랜덤 시드를 사용하였다. 다시 말해 모든 모델은 동일한 훈련 세트, 평가 세트, 테스트 세트를 사용하여 비교되었다.

통전 전후의 미세조직 분석을 위해 주사전자현미경(scanning electron microscope, SEM)이 활용되었다. 테스트 세트의 중심부에서 분석용 시편을 채취하여 연마 후 아세트산 16 mL, 피크르산 4.2 g, 증류수 10 mL, 에탄올 70 mL 혼합액에서 40초 동안 에칭하였다. SEM 분석 시 전압 및 전류는 각 10 kV, 15 μ A로 설정되었다. 소재의 결정립 크기는 linear intercept method를 기반으로 20회 측정되었다. 재결정 분율은 ImageJ 소프트웨어를 통해 측정되었다.

3. 기계 학습

3.1. Artificial Neural Network

기계 학습 기법의 하나인 인공 신경망(artificial neural network)은 회귀 문제를 다루는 유용한 도구로 알려져 있다 [17]. 인공 신경망은 일반적으로 입력층, 은닉층, 출력층으로 구성되며 각 층에는 다수의 유닛(혹은 ‘뉴런’)이 포함되어 있다. 인접 층 사이의 모든 유닛은 연결되어 있으며 각 연결의 가중치를 조절함으로써 모델을 학습시키는 구조를 가진다. 다수 유닛의 상관관계와 그 가중치를 탐색하는 인공 신경망의 구조는 다변수 비선형 문제의 해를 구하는 데 있어 높은 정확도를 보인다. 본 연구에서 다루는 통전 펄스의 최적화가 이러한 문제의 범주에 들어간다고 할 수 있다.

본 연구에 활용된 인공 신경망은 공통으로 3개의 입력층 유닛(펄스 크기, 펄스 길이, 시간) 및 1개의 출력층 유닛(온도)으로 구성되었다. 은닉층 및 유닛의 개수는 연구자가 직접 설정해야 하는 하이퍼 파라미터(hyperparameter)로 간주한다. 인공 신경망을 금속 공학에 접목한 선행 연구[18-21]를 살펴보면 1-3개의 은닉층과 1-16개의 유닛을 사용하였음이 확인된다. 이러한 구조를 모사하여, 본 연구는 다음과 같은 세 종류의 전통적인 인공 신경망을 구축하였다: ANN-1, ANN-2, ANN-3. 모델명에 기재된 숫자는 은닉층의 개수를 의미한다. 예를 들어 ANN-3 모델은 3개의 은닉층을 가진다. 상기 세 신경망의 모든 은닉층은 각 10개의 유닛으로 구성되어 있다.

전통적인 인공 신경망은 완전 연결 계층(fully connected layer)과 활성화 함수(activation function) 계층이 반복되는 다층 퍼셉트론 구조를 가진다 [17]. 완전 연결 계층 내부에 있는 하나의 유닛은 다음과 같이 수식화 가능하다:

$$\hat{y}_i = \bar{x}_i \cdot \bar{w} + b_i \quad (2)$$

이는 해당 유닛에 i 번째 샘플의 데이터($\bar{x}$)가 입력되었을 때, 가중치($\bar{w}$) 및 편향(b_i)을 적용하고 그 결과($\hat{y}_i$)를 활성화 함수 계층으로 넘겨주는 것을 의미한다. 이를 특정 층 내부의 전체 유닛으로 확장하면 식 2의 벡터 형태 대신 행렬 곱으로 한 번에 계산할 수 있다. 활성화 함수 계층의 활성화 함수는 시그모이드 함수를 사용하였다:

$$s = (1 + e^{-z})^{-1} \quad (3)$$

손실 함수(loss function)로는 평균 제곱근 편차(mean squared error, MSE)를 사용하였다:

$$MSE = \frac{1}{n} \sum_{i=1}^n (\hat{y}_i - y_i)^2 \quad (4)$$

여기서 $\hat{y}_i$ 와 y_i 는 i 번째 샘플에 대한 예측치와 실제 측정치를 각각 의미한다. 마지막으로 최적화법으로는 확률적 경사 하강법(stochastic gradient descent)을 채택하였다. 경사 하강법이란 특정 가중치에 대한 손실 함수(L)의 기울기(혹은 ‘경사’)를 구한 후, 이 값에 비례하여 가중치를 업데이트해 가는 방법이다. 예를 들어 t 번째 업데이트된 가중치를 수식으로 표현하면 다음과 같다:

$$w_t = w_{t-1} - \eta \frac{\partial L}{\partial w_{t-1}} \quad (5)$$

이 식에서 비례 계수에 해당하는 η 를 학습률(learning rate)이라 하며, 상기 세 신경망은 10^{-1} 의 학습률로 1,000 에포크(epoch) 동안 데이터를 학습하였다. 이러한 아키텍처는 가장 전통적인 인공 신경망 모델을 따른 것이다.

한편, 본 연구는 기계 학습의 예측 정확성을 더욱 개선하기 위해 심층 인공 신경망(deep neural network, DNN)을 추가로 설계하였다. DNN 구성을 위한 최소 은닉층 개수는 문헌에 따라 다소 다르지만, 일반적으로 3개 이상의 은닉층을 사용할 경우 DNN으로 간주한다 [22]. 본 연구에 도입된 DNN 모델은 단순히 은닉층 숫자를 다섯 개로 늘리는 것에 그치지 않고, 신경망 아키텍처를 구성하는 함수를 개선하고 하이퍼 파라미터를 최적화하였다. 먼저, 손실 함수를 시그모이드 함수 대신 rectifier linear unit(ReLU, [23]) 함수로 교체하였다. 은닉층 수가 증가하면 역전파 도중 미분 값이 과소화되어 기계 학습을 방해하는 문제(vanishing gradient)가 발생한다. ReLU 함수를 도입할 경우 이러한 문제를 효율적으로 방지하는 것으로 알려져 있다. 최적화법 역시 종래의 확률적 경사 하강법 대신 adaptive moment estimation(Adam, [24])으로 대체하였다. Adam 법을 간략히 수식으로 표현하면 다음과 같다.

$$w_{t+1} = w_t - \eta \frac{\hat{m}_t}{\sqrt{\hat{v}_t + \epsilon}} \quad (6a)$$

$$\hat{m}_t = \frac{m_t}{1 - \beta_1'} \quad (6b)$$

$$\hat{v}_t = \frac{v_t}{1 - \beta_2'} \quad (6c)$$

$$m_t = \beta_1 m_{t-1} + (1 - \beta_1) \frac{\partial}{\partial w_t} L(w_t) \quad (6d)$$

$$v_t = \beta_2 v_{t-1} + (1 - \beta_2) \left[\frac{\partial}{\partial w_t} L(w_t) \right]^2 \quad (6e)$$

여기서 m_t 함수는 가중치를 구할 때 이전 계산 값을 고려하여 일종의 모멘텀을 부여하는 역할을 한다. v_t 함수

Table 1. The number of units in each hidden layer of three DNN models used in this study. ‘#’ indicates the number designated to a hidden layer. ‘ n_u ’ is the number of units in the given hidden layer.

Model	#	1	2	3	4	5	6	7	8	9	10
DNN-1	n_u	2^8	2^6	2^6	2^3	2^3	2^4	2^5	2^3	2^4	-
DNN-2		2^8	2^7	2^5	2^5	2^8	2^5	2^6	-	-	-
DNN-3		2^8	2^5	2^6	2^8	2^4	2^6	2^6	2^7	2^9	-

는 지수 이동 평균(exponential moving average)을 적용하여 최근 계산 값에 더 높은 가중치를 준다. 학습률 및 기타 상수는 알고리즘 개발자가 제시한 기본값[24]을 사용하였다: $\eta = 10^{-3}$, $\beta_1 = 0.900$, $\beta_2 = 0.999$, $\varepsilon = 10^{-8}$. 반면 과적합을 방지하기 위해 자주 사용되는 early stop routine 및 dropout routine[25]은 예비 실험 단계에서 예측 정확성을 오히려 떨어뜨리는 경향을 보였기에 본 연구에서는 배제되었다.

DNN 모델은 10^4 - 10^5 단위의 거대한 파라미터 규모를 가지므로 격자 탐색을 통해 은닉층 및 유닛 개수를 최적화하려면 막대한 컴퓨터 자원이 소모된다. 따라서 본 연구는 하이퍼 파라미터 최적화를 위해 hyperband 알고리즘[26]을 도입하였다. Hyperband 알고리즘의 원리를 간략하게 설명하면 다음과 같다. 우선 주어진 하이퍼 파라미터 탐색 범위 내에서 임의의 설정을 추출한다. 해당 하이퍼 파라미터를 적용하여 짧은 에포크 학습 후 예측 성능을 비교한다. 성능이 낮은 하이퍼 파라미터는 배제하고, 남은 값으로 조금 더 긴 에포크 학습 후 재차 성능을 비교한다. 마지막 하이퍼 파라미터가 남을 때까지 이 과정을 반복하여 최적 조건을 구한다. 본 연구는 m 개의 은닉층($5 \leq m \leq 10$ 의 정수) 및 각 은닉층 당 2^n 개의 유닛($3 \leq n \leq 9$ 의 정수)으로 탐색 범위를 설정하였다. Hyperband 알고리즘을 활용하여 90개의 DNN 모델 중 정확도 기준 상위 3개를 선정하여 각각 DNN-1, DNN-2, DNN-3으로 명명하였다. Table 1에 이들 DNN 모델의 하이퍼 파라미터를 정리하였다.

3.2. Extreme Gradient Boosting (XGBoost)

XGBoost는 2016년 Chen과 Guestrin[27]이 제시한 기계 학습 기법으로, 최근 수년간 Kaggle Competition이나 KDD Cup 등 세계적인 인공지능 경진대회에서 우수한 성적을 기록하며 주목받고 있다. XGBoost는 기본적으로 gradient tree boosting 기법의 연산 속도를 개선한 알고리즘이다. Gradient tree boosting은 다수의 의사 결정 나무(decision tree)를 엮어서 예측 정확성을 향상한 앙상블 모델이다. 이 기법에서 i 번째 샘플에 대한 예측치는 다음과 같이 수식화할 수 있다:

$$\hat{y}_i = \sum_{k=1}^K f_k(x_i) \quad (7)$$

f_k 는 함수 공간의 벡터이며 K 는 의사 결정 나무의 개수이다. Tree boosting 모델의 목표는 목적 함수(ϕ)를 최소화하는 f_k 의 집합을 찾는 것이다. 여기서 목적 함수는 손실 함수에 과적합을 방지하기 위한 규제(Ω)를 추가한 것으로 다음과 같은 수식으로 표현된다:

$$\phi = \sum_i L(y_i, \hat{y}_i) + \sum_k \Omega(f_k) \quad (8)$$

식 8은 함수를 변수로 가지므로 통상적인 방법 대신 반복(iteration)을 통해 최솟값을 구한다. t 번째 업데이트된 목적 함수는 식 9a와 같이 표현되며, 이는 식 9b로 근사할 수 있다.

$$\phi_t = \sum_i L(y_i, \hat{y}_{i,t-1} + f_t(x_i)) + \Omega(f_t) \quad (9a)$$

$$\tilde{\phi} = -\sum_j \frac{G_j}{H_j + \lambda} + \gamma T \quad (9b)$$

상기 식에서 T 는 주어진 의사 결정 나무의 출력 차원(혹은 ‘잎 수’)이다. G_j 는 이 모델의 j 번째 그룹에 속한 샘플에 대한 손실 함수의 1차 기울기 총합이다. H_j 는 동일한 손실 함수의 2차 기울기 총합이다. γ 와 λ 는 규제의 강도를 정하는 상수이다. 식 9a부터 식 9b까지의 더욱 상세한 유도는 선행 연구논문[27]에 기술되어 있다. 요약하자면, gradient tree boosting의 핵심은 식 9b를 최소화하는 모델을 탐색하여 예측 정확성을 높이는 데 있다. 본 연구에서 사용한 XGBoost 기반 첫 모델은 이 알고리즘의 하이퍼 파라미터 기본값[27]을 그대로 사용하여 설계하였으며 이를 XGB-1 모델로 명명하였다. 향후 ‘결과 및 고찰’ 항목에서 이 모델의 개선 방법에 대해 더욱 상세히 설명할 것이다. 본 연구에 사용된 모든 코드는 Python ver. 3.6으로 작성되었다.

4. 결과 및 고찰

탄성 영역에서 소재의 열응력(σ_T)은 온도 변화(ΔT)와

Fig. 1. Temperature variation of Mg alloys subjected to the single electropulse with various pulse magnitudes and durations. The red curve corresponds to the test dataset obtained at a current density of 59.5 A/mm² and application time of 1,700 ms.

다음과 같은 관계에 있다.

$$\sigma_T = E\alpha\Delta T \quad (10)$$

연구 소재의 탄성계수(E)는 45 GPa, 열팽창계수(α)는 27.0 $\mu\text{m}/(\text{m}^\circ\text{C})$ 이다 [28]. 식 10에 의하면 205.8 $^\circ\text{C}$ 이상의 온도 변화가 야기되었을 때 열응력이 소재의 압축 항복 강도(250 MPa, [28])를 초과한다. 통전 시험 전 판상 시편의 초기 온도는 15.4–22.6 $^\circ\text{C}$ 로 관측되었으므로, 이론적으로 판상 시편 온도가 221.2 $^\circ\text{C}$ 를 초과하지 않는 이상 열응력에 의한 소성 변형을 방지할 수 있을 것으로 판단하였다. 이러한 가정에 기반하여 최고 온도가 193.5–212.4 $^\circ\text{C}$ 범위에 들어오도록 전류 펄스의 크기 및 길이를 조정하였다. 실제로 통전 시험 종료 후 확인한 결과 어떤 시편에서도 열응력에 따른 뒤틀림이나 굽힘의 증거는 발견되지 않았다.

Fig. 1에 통전 조건에 따른 온도 곡선을 도시하였다. 모든 시편은 최대 4초 이내에 급격히 가열되어 최고 온도에 도달 후 상대적으로 완만한 속도로 초기 온도 영역까지 냉각되는 모습을 보인다. 모든 온도 곡선은 비선형 가열·냉각 경향을 드러내며, 이는 통전 가열의 온도 예측 난도를 높이는 원인이 된다. 일부 온도 곡선의 경우 냉각 시 순간적인 온도의 급락이 관측되는데, 이는 적외선 카메라의 일시적 지연에 기인한 측정 오류인 것으로 판단된다. 본 연구에서는 이러한 실험 오류가 기계 학습 정확성에 유의미한 악영향을 미치는지 판단하고자 급락 값을 제거하지 않고 그대로 사용하였다.

Fig. 2에 테스트 세트 소재의 SEM 분석 결과를 나타내었다. 본 연구에 사용된 압출재는 (i) 압출 방향으로 연신

Fig. 2. SEM micrograph of sample subjected to the electropulse at a current density of 59.5 A/mm² and application time of 1,700 ms (i.e., the test dataset). “ED” indicates the extrusion direction.

된 조대한 미세결정 결정립 및 (ii) 등축상의 미세한 재결정 결정립으로 구성되어 있다. 재결정된 결정립의 평균 크기는 5.59 μm 이며 그 분율은 52.1%로 측정되었다 [29]. 동일한 방법으로 측정한 테스트 세트 소재의 평균 결정립 크기는 8.59 μm , 재결정 분율은 58.7 %였다. 이를 통해 전류 펄스 인가에 따른 소재의 재결정이 야기되었음을 알 수 있다. 이는 마그네슘 합금의 통전에 대한 선행 연구와도 일치하는 결과이다. 예를 들어 Li 등[30]은 AZ31 냉연재에 33 A/mm² 크기의 전류 펄스를 인가하여 270 $^\circ\text{C}$ 에서 7초만에 재결정이 완료되었음을 보고하였다. Jin 등[31]은 ZK60 마그네슘 합금에 4,500 A/mm² 크기의 고전류 펄스를 인가하여 통상적인 정적 재결정 온도(300 $^\circ\text{C}$)보다 훨씬 낮은 온도(132 $^\circ\text{C}$)에서 재결정을 유도하였음을 보고하였다.

전류 펄스 형태에 따른 온도 변화를 예측하는 기존 방식은 물리 함수를 기반으로 한 접근법이다. 전류 펄스에 의한 전기에너지가 일정 비율(ζ)로 줄의 법칙에 따른 열량으로 전환된다고 가정 시, 줄열에 따른 온도 변화는 다음과 같이 표현할 수 있다:

$$\Delta T = \frac{\zeta \rho_c J^2 \Delta t}{c \rho} \quad (11)$$

여기서 J 와 Δt 는 전류 펄스의 크기(전류 밀도) 및 길이(적용 시간)를 의미한다. 또한 본 연구에 사용된 마그네슘 합금의 비열(c)은 1.00 J/g $^\circ\text{C}$, 밀도(ρ)는 1.83 g/cm³, 비저항(ρ_c)은 6.00 $\times 10^{-8}$ $\Omega\cdot\text{m}$ 이다 [28]. Fig. 3는 식 11을 바탕으로 계산한 최대 가열 온도를 측정치와 비교하여 나타낸 것이다. 즉, 에너지 손실이 없다는 가정 하의 줄열 예측치를 실제 측정치와 비교하였다. 이 도표의 기울기로

Fig. 3. Experimental peak temperatures arranged in terms of Eq. 11.

부터 $\zeta = 0.89$ 가 산출되었다. 그러나 이 비례 상수의 물리적 의미가 전기에너지와 열에너지 사이의 전환 효율임을 고려하면 0.89는 과하게 높은 수치로 보이며, 이를 통해 식 11의 불완전성이 간접적으로 드러난다. 한편 식 11을 통해 예측한 테스트 세트의 최고 온도는 198.6 °C로, 측정치 대비 8.7 °C의 오차를 보인다. 이 값은 오차범위 4.2 %의 비교적 준수한 예측치이지만, 본 연구에서는 기계 학습을 통해 더욱더 높은 예측 정확성을 확보하고자 하였다. 또 한 가지 주목할 점은 식 11과 같은 접근법이 전류 펄스 인가에 의한 가열 단계에만 적용된다는 점이다. 펄스 인가 종료 후의 냉각 단계에서는 별도의 물리 함수가 요구되므로 복수의 회귀 분석이 강제되는 문제가 있다. 본 연구에서는 이러한 문제 역시 기계 학습을 도입하여 모델을 단일화함으로써 해결하고자 하였다.

Fig. 4은 전통적인 ANN 모델로 테스트 세트의 온도 곡선을 예측한 후 이를 실제 측정치와 비교하여 나타낸 것이다. 기계 학습 모델의 신뢰성을 정량적으로 검증하기 위해 평균 제곱근 오차(root mean square error, RMSE) 척도가 도입되었다. 이는 식 4의 MSE 척도에 제곱근을 적용한 수치이다. ANN-1, ANN-2, ANN-3 모델의 RMSE는 각각 0.118 °C, 0.078 °C, 0.055 °C로 측정되었다. 이는 은닉층 수가 증가함에 따라 모델 정확성이 급격히 개선되었음을 나타낸다.

세 인공 신경망 모두 펄스 인가 직후(1초 이전)의 급격한 온도 상승과 실험 후반부(8초 이후)의 완만한 온도 하강은 대체로 잘 예측하였다. 상기 언급한 소수의 급락 값은 예측 정확성에 큰 영향을 미치지 못한 것으로 판단된다. 그러나 전류 펄스가 종료되며 가열에서 냉각으로 전환되는 구간(이하 ‘가열-냉각 경계’)에서는 모델 별로 뚜렷한 차이가 드러난다. 이 구간에서는 시편이 최고 온도에 도달 후 급격히 냉각되며 온도가 비선형적으로 변화하게 된다. 따라서 온도 예측 난도가 가장 높고 기계 학습 모델의 예측 정확성을 평가하기 용이하다. Fig. 4a 내부의 확대도를 보면, ANN의 은닉층 수가 증가함에 따라 최고 온도 예측치가 실제 측정치에 점차 근접해 가는 모습이 명백히 관측된다. ANN-1 모델의 경우 최고 온도 예측에서 15.3 °C의 오차를, 최고 온도 도달 시간 예측에서 730 ms의 오차를 보였다. 이는 본 연구에서 검증한 모든 예측 모델 중 가장 큰 오차이다. 실제로 Fig. 4a에 그려진 ANN-1의 예측 곡선은 타 모델과 비교해 육안으로도 확인 가능할 정도로 뚜렷한 불일치를 드러내고 있다. ANN-2의 경우 오차가 약 절반으로 감소하여(각 8.7 °C 및 450 ms) 식 11에 나타낸

Fig. 4. ANN prediction of (a) temperature variation and (b) temperature gradient. The inset in Fig. 4a is a magnified image focusing on the heating-cooling boundary. The dotted line in Fig. 4b indicates a rapid reduction of temperature gradient for the experimental data. The prediction of DNN-2 model is also included in Fig. 4b for further discussion.

물리 함수 기반 모델과 유사한 수준의 예측 정확성을 보였다. ANN-3 모델은 정확성이 더욱 개선되어 각각 4.8 °C 및 240 ms의 오차를 기록하였다. 이는 인공 신경망 모델에 은닉층을 추가함으로써 모델 복잡도가 증가한 까닭으로 판단된다.

상술한 바와 같이 ANN-3 모델의 경우 물리 기반 함수 및 나머지 두 인공 신경망에 비해 크게 개선된 예측 정확성을 보였다. 그러나 이 모델이 산출한 오차를 측정치에 대한 비율로 살펴볼 경우, 온도 및 시간 예측에 있어 서로 다른 양상이 나타난다. ANN-3 모델이 예측한 최고 온도는 측정치 대비 2.3%의 절대 비율 오차(absolute percentage error)를 보인다. 이를 통해 이 모델이 온도 예측에 있어서 우수한 정확성을 가지고 있음을 재차 확인할 수 있다. 반면, 최고 온도 도달 시간을 살펴보면 ANN-3 모델은 상대적으로 큰 절대 비율 오차(14.3%)를 기록하였다. 이는 높은 에너지 효율을 보이는 통전 가열 특성상, 최대 온도에도 도달하는 시점이 극단적으로 짧았기 때문이다. 또 한 가지 주목할 점은 가열-냉각 경계를 전후로 온도 곡선이 보이는 형태이다. ANN-3 모델을 포함한 전통적 인공 신경망은 최대 온도를 중심으로 하여 상대적으로 대칭을 이루는 형태의 온도 곡선을 예측하였다. 그러나 이는 가열-냉각 경계에서 실제 관측된 비선형적 온도 변화에 배치되는 예측이다. 이러한 차이는 측정치와 ANN 예측치의 온도 변화율, 즉 온도 곡선의 1차 미분 값을 비교할 시 더욱 뚜렷하게 관측된다. Fig. 4b에 점선으로 표시된 바와 같이, 실제 온도 변화율은 가열-냉각 경계를 지나는 순간 급감하는 경향을 보인다. 대조적으로 ANN-1, ANN-2, ANN-3 모델이 예측한 온도 변화율은 모두 완만하고 연속적인 감소를 나타내고 있다.

Fig. 5에 각 DNN 모델이 산출한 온도 예측치를 측정치와 함께 도시하였다. 1,000 에폭 학습 이후 DNN-1, DNN-2, DNN-3 모델의 RMSE는 각각 0.059 °C, 0.037 °C, 0.037 °C로 측정되었다. 이는 전통적 인공 신경망이 산출한 오차에 비하여 압도적으로 작은 수치이다. 각 DNN 모델이 예측한 최고 온도는 측정치 대비 각각 3.1 °C, 2.8 °C, 4.7 °C의 오차를 기록하였다. 최고 온도 도달 시간은 각각 182 ms, 178 ms, 178 ms의 오차를 보였는데, 이는 절대 비율 오차 기준으로 8% 미만의 매우 작은 값이다. 상기 모든 결과는 최적화된 DNN 아키텍처가 예측 정확성을 혁신적으로 개선하였음을 보여준다. DNN 모델의 우수한 예측 성능은 앞서 논의된 온도 변화율 관점에서도 재차 확인할 수 있다. 가장 우수한 성능을 보인 DNN-2 모델을 활용하여 온도 변화율을 산출한 후 Fig. 4b에 도시하였다.

Fig. 5. DNN prediction of temperature variation. The inset is a magnified image focusing on the heating-cooling boundary.

Fig. 6. RMSE values calculated from 792 combinations of XGBoost models with various learning rates and maximum depths. The arrows indicate the selection of XGB-2 and XGB-3 models with the lowest RMSE values.

이 모델은 실제 시편이 가열-냉각 경계에서 보인 급격한 온도 변화율 감소를 잘 묘사하고 있다. 이러한 경향은 전통적 인공 신경망을 따라 구축한 ANN 모델과 명백히 대조된다.

XGBoost 알고리즘 역시 인공 신경망 못지않은 다양한 하이퍼 파라미터를 가지고 있다. 본 연구에서는 그중 학습률과 의사 결정 나무의 최대 깊이(이하 ‘최대 깊이’)를 변경하여 기계 학습 모델의 성능을 향상하고자 하였다. 학습률은 0.01부터 0.99까지 0.01 단위로, 최대 깊이는 3부터 10까지 1 단위로 탐색 범위를 설정하였다. 이 탐색 범위에서 가능한 조합은 총 792개이며, 모든 조합에 대해 격자 탐색을 자동으로 수행하는 코드를 작성하였다. Fig. 6에 792개의 XGBoost 모델이 산출한 RMSE를 정리하였다. 학

Fig. 7. (a) XGBoost prediction of temperature variation and (b) comparison of the best models among the investigated algorithms. The inset in Fig. 7a is a magnified image focusing on the heating-cooling boundary.

습률과 최대 깊이가 동시에 낮은 모델은 대체로 높은 오차를 보인다. 그러나 전체 모델에 대해 탐색 변인과 RMSE 간 정량적 관계식을 찾는 것은 어려워 보인다. 대신, 본 연구는 격자 탐색 결과로부터 RMSE가 현저히 낮은 두 개의 특이점을 발견하였다. 각 특이점을 구성하는 하이퍼 파라미터를 사용하여 XGB-2 및 XGB-3 모델을 설계하였다. 구체적으로 XGB-2는 학습률 0.04 및 최대 깊이 10으로, XGB-3는 학습률 0.94 및 최대 깊이 4로 설계되었다.

Fig. 7에 XGBoost 알고리즘을 통해 예측한 온도 곡선을 도시하였다. 각 XGB 모델이 산출한 시계열 온도 예측치는 실험 측정치에 비해 높은 불규칙성을 보인다. 이는 ANN 모델(Fig. 4a) 및 DNN 모델(Fig. 5)이 예측한 연속적인 온도 변화 추세와 대비되는 경향이다. XGB-1 모델의 경우 가열-냉각 경계의 급격한 온도 변화를 전혀 예측하지 못하고 있음을 육안으로 확인할 수 있다. 이 모델의 RMSE(0.105 °C) 역시 ANN-1 모델과 더불어 본 연구에서 가장 낮은 예측 정확성을 보여준다. 상술한 바와 같이 XGB-1 모델은 하이퍼 파라미터 기본값을 그대로 사용하여 설계되었다. 그러므로 XGBoost 알고리즘을 활용하기 위해서는 하이퍼 파라미터 최적화가 반드시 선행되어야 함을 추론할 수 있다.

실제로 격자 탐색을 통해 하이퍼 파라미터를 최적화시킨 XGB-2와 XGB-3 모델의 경우 예측 정확성이 크게 개선되었음이 확인된다. XGB-2와 XGB-3 모델이 산출한 RMSE는 각각 0.078 °C 및 0.076 °C로, 이는 XGB-1의 RMSE에 비해 크게 낮은 수치이다. 특히 XGB-2 모델의 경우 가열-냉각 경계에서 매우 우수한 예측 정확성을 보였다. 이 모델이 예측한 최고 온도 및 최고 온도 도달 시간은 측정

치 대비 0.3 °C와 40 ms의 오차를 기록하였다. 다만 XGB 모델의 전체 시계열에 대한 전반적인 예측 정확성은 인공 신경망에 비해 낮은 편이다. Fig. 7b에 나타난 바와 같이 XGB-2 모델은 가열 구간에서 과다 예측을 보이며, 오차가 가열-냉각 경계 이전까지 점차 증가하다 냉각 구간에 들어서며 확연히 감소한다. 반면 ANN-3 및 DNN-2 모델은 상대적으로 개선된 예측 정확성을 보이며, 특히 후자의 경우 전 시계열에서 균일하게 작은 오차를 나타낸다. 따라서 예측 정확성 측면에서는 심층 인공 신경망 아키텍처 및 하이퍼 파라미터 최적화를 통해 설계한 DNN-2 모델이 가장 우수하다고 결론지을 수 있다.

마지막으로 예측 정확성이 아닌 학습 시간 관점에서 각 기계 학습 모델을 평가해보고자 한다. 본 연구에 활용된 9가지 모델의 학습 시간을 Fig. 8에 나타내었다. 로그 스케일로 학습 시간을 비교해 보면 XGB 모델의 학습 속도가 ANN 및 DNN 모델과 비교해 압도적으로 빠르다는 사실이 뚜렷하게 드러난다. 각 그룹 내에서 정확도가 가장 높은 모델을 뽑아 학습 속도를 비교해 보면 그 차이는 더욱 명백하다. ANN-3 모델의 경우 학습에 약 7분가량(408,241 ms) 소요되었다. DNN-2 모델의 경우 본 연구에서 평가된 모든 기계 학습 모델 중 가장 높은 정확성을 보이나, DNN-3의 뒤를 이어 두 번째로 긴 학습 시간(840,701 ms)이 소요되었다. 이는 ANN-2 모델의 학습 시간보다 두 배 이상 큰 값이다. 상기 두 모델과 대조적으로, XGB-2 모델의 경우 준수한 예측 정확성을 보였음에도 불구하고 학습에 소요된 시간은 1초 미만(784 ms)에 불과하였다. 이는 DNN-2 모델 대비 1,000배 이상 빠른 속도이며, XGBoost 알고리즘의 하이퍼 파라미터 최적화에 격자

Fig. 8. Comparison of log-scale learning time among the nine machine-learning models used in this study.

탐색을 적용할 수 있었던 이유이기도 하다. 단일 펄스의 통전 가열에 집중된 본 연구는 상대적으로 적은 데이터 및 단순한 변수를 가진 문제를 다루고 있어 DNN 모델의 긴 학습 시간이 큰 단점으로 부각되지 않았다. 그러나 향후 본 연구가 빅데이터-다변수 환경으로 확장되면 XGBoost 알고리즘이 대안으로 활용될 여지가 클 것으로 기대된다. 다시 말해 (1) 다양한 소재에 (2) 다중 펄스를 인가하는 (3) 통전 성형을 해석하고자 할 경우 필요한 컴퓨터 자원이 급격히 증가하게 되므로, 이에 따라 XGBoost 알고리즘이 더욱 적극적으로 활용될 수 있으리라 예측된다.

5. 결 론

본 연구는 인공 신경망 및 XGBoost 알고리즘을 기반으로 한 9개의 기계 학습 모델을 활용하여 전류 펄스에 의해 야기되는 비선형 온도 변화를 예측하였다. 본 연구의 결과는 다음과 같이 요약된다.

ANN 모델의 경우 은닉층 수에 비례하여 예측 정확성이 증가하였다. 대표 모델인 ANN-3는 최고 온도를 비교적 잘 예측하였으나 최고 온도 도달 시간에서 다소 큰 절대 비율 오차를 기록하였다. 더욱이 실험에서 관측된 가열-냉각 경계의 급격한 온도 변화가 이들 모델에서는 제대로 묘사되지 않았다.

DNN 모델은 RMSE 및 절대 비율 오차 측도 양면에서 정확성이 크게 개선되었으며, 가열-냉각 경계의 급격한 온도 변화를 감소 역시 잘 예측하였다. 대표 모델인 DNN-2는 본 연구에서 시험한 9개의 기계 학습 모델 중 가장 우

수한 예측 정확성을 보였다.

XGB 모델은 하이퍼 파라미터 기본값을 적용 시 열악한 예측 정확성을 보였으며, 이를 통해 XGBoost 알고리즘에는 하이퍼 파라미터 최적화가 필수적임을 확인하였다. 대표 모델인 XGB-2는 하이퍼 파라미터 최적화 이후 준수한 예측 정확성을 보였을 뿐 아니라, 인공 신경망 아키텍처 대비 10²-10³배 빠른 학습 속도를 드러냈다. 이는 향후 빅데이터-다변수 문제를 해결하는 데 큰 강점으로 작용할 것으로 기대된다.

ACKNOWLEDGEMENT

This work was supported by a 2-Year Research Grant of Pusan National University.

REFERENCES

1. M. Mesbah, A. Fattahi, A. R. Bushroa, G. Faraji, K. Y. Wong, W. J. Basirun, A. Fallahpour, and B. Nasiri-Tabrizi, *Met. Mater. Int.* (2019). <https://doi.org/10.1007/s12540-019-00495-w>
2. Y. B. Chun and C. H. J. Davies, *Metall. Mater. Trans. A* **42**, 4113 (2011).
3. T. T. T. Trang, J. H. Zhang, J. H. Kim, A. Zargar, J. H. Hwang, B. C. Suh, and N. J. Kim, *Nat. Comm.* **9** (2018).
4. H.-D. Nguyen-Tran, H.-S. Oh, S.-T. Hong, H. N. Han, J. Cao, S.-H. Ahn, and D.-M. Chun, *Int. J. Precis. Eng. Man. - GT* **2**, 365 (2015).
5. J. Magargee, F. Morestin, and J. Cao, *J. Eng. Mater. Technol.* **135**, 41003 (2013).
6. T. Lee, J. Magargee, M. K. Ng, and J. Cao, *Int. J. Plast.* **94**, 44 (2017).
7. C. D. Ross, D. B. Irvin, and J. T. Roth, *J. Eng. Mater. Technol.* **129**, 342 (2007).
8. T. A. Perkins, T. J. Kronenberger, and J. T. Roth, *J. Manuf. Sci. Eng.* **129**, 84 (2006).
9. H.-J. Jeong, M.-J. Kim, J.-W. Park, C. D. Yim, J. J. Kim, O. D. Kwon, P. P. Madakashira, and H. N. Han, *Mater. Sci. Eng. A* **684**, 668 (2017).
10. H.-J. Jeong, J. Park, K. J. Jeong, N. M. Hwang, S.-T. Hong, and H. N. Han, *Int. J. Precis. Eng. Man. - GT* **6**, 315 (2019).
11. W. A. Salandro, J. J. Jones, T. A. McNeal, J. T. Roth, S.-T. Hong, and M. T. Smith, *J. Manuf. Sci. Eng.* **132**, 51016 (2010).
12. J. Lee, H. J. Bong, Y.-S. Lee, D. Kim, and M.-G. Lee, *Metall.*

- Mater. Trans. A* **50**, 2720 (2019).
13. D. R. Cassar, A. C. P. L. F. de Carvalho, and E. D. Zanotto, *Acta Mater.* **159**, 249 (2018).
 14. T. Lee, B. J. Kwak, J. Yu, J. H. Lee, Y. Noh, and Y. H. Moon, *Mater. Lett.* **269**, 127652 (2020).
 15. S.-J. Jeong, I.-K. Hwang, I.-S. Cho, and H.-S. Kim, *Korean J. Met. Mater.* **57**, 184 (2019).
 16. A. Agrawal and A. Choudhary, *APL Mater.* **4**, 53208 (2016).
 17. D. E. Rumelhart and J. L. McClelland, *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*, MIT Press, Cambridge (1987).
 18. C. H. Park, D. Cha, M. Kim, N. S. Reddy, and J.-T. Yeom, *Met. Mater. Int.* **25**, 768 (2019).
 19. S. Zhu, J. Li, L. Ma, C. He, E. Liu, F. He, C. Shi, and N. Zhao, *Mater. Lett.* **233**, 294 (2018).
 20. S. Gangi Setti and R. N. Rao, *Rare Metals* **33**, 249 (2014).
 21. C.-L. Li, P. L. Narayana, N. S. Reddy, S.-W. Choi, J.-T. Yeom, J.-K. Hong, and C. H. Park, *J. Mater. Sci. Technol.* **35**, 907 (2019).
 22. Y. LeCun, Y. Bengio, and G. Hinton, *Nature* **521**, 436 (2015).
 23. V. Nair and G. E. Hinton, *ICML 2010 – Proceeding of the 27th International Conference on Machine Learning*, pp. 807-814, ICML 2010, Haifa, Israel (2010).
 24. D. P. Kingma and J. L. Ba, *3rd International Conference on Learning Representations - Conference Track Proceedings*, ICLR 2015, San Diego, USA (2015).
 25. N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, *J. Mach. Learn. Res.* **15**, 1929 (2014).
 26. L. Li, K. Jamieson, G. DeSalvo, A. Rostamizadeh, and A. Talwalkar, *J. Mach. Learn. Res.* **18**, 1 (2018).
 27. T. Chen and C. Guestrin, *Proceedings of ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 785-794, KDD 2016, San Francisco, USA (2016).
 28. M. M. Avedesian and H. Baker, *ASM Specialty Handbook: Magnesium and Magnesium Alloys*, ASM International, Cleveland (1999).
 29. T. Lee, B. J. Kwak, T. Kong, J. H. Lee, S. W. Lee, and S. H. Park, *J. Alloy. Compd.* **803**, 434 (2019).
 30. X. Li, X. Li, J. Zhu, X. Ye, and G. Tang, *Scr. Mater.* **112**, 23 (2016).
 31. W. Jin, J. Fan, H. Zhang, Y. Liu, H. Dong, and B. Xu, *J. Alloy. Compd.* **646**, 1 (2015).