

ORIGINAL ARTICLE

LSTM 순환 신경망을 이용한 재료의 단축하중 하에서의 응력-변형률 곡선 예측 연구

변훈, 송재준*

서울대학교 공과대학 에너지시스템공학부

Prediction of the Stress-Strain Curve of Materials under Uniaxial Compression by Using LSTM Recurrent Neural Network

Hoon Byun and Jae-Joon Song*

Department of Energy Systems Engineering, Seoul National University

*Corresponding author: songjj@snu.ac.kr

Received: June 11, 2018

Revised: June 17, 2018

Accepted: June 26, 2018

ABSTRACT

LSTM (Long Short-Term Memory) algorithm which is a kind of recurrent neural network was used to establish a model to predict the stress-strain curve of an material under uniaxial compression. The model was established from the stress-strain data from uniaxial compression tests of silica-gypsum specimens. After training the model, it can predict the behavior of the material up to the failure state by using an early stage of stress-strain curve whose stress is very low. Because the LSTM neural network predict a value by using the previous state of data and proceed forward step by step, a higher error was found at the prediction of higher stress state due to the accumulation of error. However, this model generally predict the stress-strain curve with high accuracy. The accuracy of both LSTM and tangential prediction models increased with increased length of input data, while a difference in performance between them decreased as the amount of input data increased. LSTM model showed relatively superior performance to the tangential prediction when only few input data was given, which enhanced the necessity for application of the model.

Keywords: Recurrent neural network, LSTM, Uniaxial compression test, Stress-strain curve

초록

이 논문에서는 재료의 단축하중 하에서의 응력-변형률 곡선을 예측하기 위하여 순환 신경망의 일종인 LSTM(Long Short-Term Memory) 알고리즘을 사용하였다. 석고와 규사를 혼합해 만든 재료에 일축압축시험을 수행하여 얻은 응력-변형률 데이터를 이용하였으며, 낮은 응력 구간의 초반 데이터를 활용해서 파괴 전까지의 거동을 예측하였다. 앞부분의 데이터를 활용하여 단계적으로 뒤쪽 구간의 값을 예측하는 LSTM 순환 신경망의 구조상 큰 응력에 대응하는 변형률을 예측할 경우에는 앞쪽 구간의 오차가 누적되어 실측값과 차이가 늘어났으나 전반적으로 높은 정확도로 응력-변형률 곡선을 예측하였다. 예측에 사용한 초기 데이터의 길이가 늘어나는 경우 정확도는 조금 증가했다. 그러나 접선을 이용한 단순 예측과의 성능 차이는 초기 데이터의 길이가 작은 경우에 두드러졌으며, 적은양의 데이터로도 응력-변형률 곡선 전체 구간의 예측을 가능하게 한다는 점으로부터 신경망 모델의 필요성을 확인하였다.

핵심어: 순환 신경망, LSTM, 일축압축시험, 응력-변형률 곡선



1. 서론

점탄성을 보이는 재료는 단축하중이 가해졌을 경우 파괴 전까지 차례대로 초기 비선형 구간, 탄성 거동 구간, 파괴 직전의 비선형 구간을 보인다. 이러한 점탄성 재료의 거동을 모사하기 위해 스프링, dashpot, slider의 조합으로 이루어진 유변학적 모델이 사용되어 왔다(Hudson and Harrison, 2000). 대표적으로 스프링과 dashpot을 직렬로 연결한 Maxwell 모형, 병렬로 연결한 Kelvin-Voigt 모형, 그리고 두 모형을 직렬로 연결한 Burger 모형이 있으며, Burger 모형은 Creep 해석에 주로 사용된다(Hardy, 1967). 이후에도 다양한 연구자들이 스프링, dashpot, slider를 조합하여 점탄성 거동을 나타내는 모델을 수립하였다(Guan et al., 2008, Karami and Fahimifa, 2013, Li et al., 2014, Zhao et al., 2016). 기존에 사용되어온 모델들은 실험으로 얻어진 응력-변형률 데이터를 사용하여 반복 계산을 통해 각 모델에서의 매개변수값들을 찾아낼 수 있으나(Arindam and Prabir, 2012), 일부분의 데이터로 전체 구간의 곡선을 모델링하는 경우에는 정확성 면에서 한계가 있을 수 밖에 없다.

본 연구에서는 점탄성 모델이 일축압축시험 초반부 응력-변형률 데이터만을 사용하여 후반부의 응력-변형률 곡선을 예측하는 것에 한계가 있다는 것을 확인한 후, 신경망 모델을 도입하여 이와는 차별화된 결과를 얻었다. 결과적으로 이론적인 모델 도입 없이 오직 응력-변형률 데이터만을 사용하여 학습 과정을 거친 후 적은 양의 초기 단계의 응력-변형률 데이터로부터 향후의 거동을 예측하였다. 사용한 알고리즘은 순환 신경망(Recurrent Neural Network, RNN)의 한 종류인 LSTM(Long Short-Term Memory)이며, 입력 구간 이상의 영역에서 임의의 응력 수준에서의 변형률 값이 계산되도록 하였다. 이 모델은 오직 응력-변형률 데이터만으로 만들어진 모델이므로, 점탄성 모형과 달리 응력-변형률 곡선이 특정 지배방정식을 만족토록 제한하지 않아 더욱 다양한 형태의 곡선 이해와 도출이 가능하다. 따라서 신경망을 통한 예측 방법은 기존 방식에 비해 표현력이 더욱 좋으므로 높은 정확도로 주어진 데이터에 맞는 모델을 만들 수 있다. 결과적으로 신경망 모델을 사용하면 초기 단계의 응력-변형률 곡선만으로 이후 곡선 예측이 가능하게 되어, 가압에 따른 시료 손상을 최소화하면서 파괴 전까지의 응력-변형률 곡선을 얻을 수 있다는 점에서 의의가 있다. 모델 학습에 사용한 데이터는 석고와 규사를 혼합하여 제작한 재료에 일축압축시험을 수행하여 얻어진 응력-변형률 값들이며, 초기의 응력-변형률 값을 활용하여 재료의 파괴 전까지의 거동을 예측해내었다.

2. 모델 구성 및 해석

2.1 점탄성 모델

점탄성 모델 중 Maxwell 모형과 Kelvin-Voigt 모형은 단순하여 변곡점이 존재하는 곡선을 표현하는 것에 한계가 있기 때문에 Burger 모형을 실험에 사용하였다. 식 (1)은 Burger 모형의 방정식이다(Skrzypek and Ganczarski, 2015).

$$\frac{k_1 k_2}{E_2} \ddot{\epsilon} + k_1 \dot{\epsilon} = \frac{k_1 k_2}{E_1 E_2} \ddot{\sigma} + \left(\frac{k_1}{E_1} + \frac{k_1}{E_2} + \frac{k_2}{E_2} \right) \dot{\sigma} + \sigma \quad (1)$$

식 (1)의 k_1, k_2, E_1, E_2 는 Burger 모형에의 매개변수들이다. 본 연구에서는 일정한 속도(0.4 mm/min)로 변형률을 증가시키며 실험하였다. 따라서 $\dot{\epsilon}$ 는 상수이며, $\ddot{\epsilon} = 0$ 이 된다. 그러나 실측값은 실험 장비의 미세한 제어 오차로 인해 $\dot{\epsilon}$ 값이 실험 과정 중 미세

하게 진동한다. 따라서 측정된 $\dot{\epsilon}$ 와 $\ddot{\epsilon}$ 값을 계산에 사용하였다. 주어진 실험 데이터를 사용해 식 (1)의 각 계수들을 예측하는 과정을 수행하였다. 변수 계산 과정은 Python과 Tensorflow 패키지를 활용하였으며, 초기값을 무작위로 설정한 뒤, 식 (2)에 표시한 오차 함수를 활용하여, 경사하강법을 통해 Burger 모델의 변수 최적화를 진행하였다.

$$loss\ function = \left[\frac{k_1 k_2}{E_2} \ddot{\epsilon} + k_1 \dot{\epsilon} - \frac{k_1 k_2}{E_1 E_2} \ddot{\sigma} - \left(\frac{k_1}{E_1} + \frac{k_1}{E_2} + \frac{k_2}{E_2} \right) \dot{\sigma} - \sigma \right]^2 \quad (2)$$

최적화 과정을 통해 매개변수를 결정한 후, 식 (1)에 이 값들을 대입하여 미분방정식을 풀고 응력-변형률 곡선을 산출했다. 실제로 $\dot{\epsilon}$ 값은 매우 작은 변화를 보이기 때문에 방정식을 푸는 과정에서는 이 값을 상수로 놓고 계산했다. 계산된 일반해의 경계조건은 변형률의 양끝값을 사용하였다.

2.2 순환 신경망

이 연구에는 인공신경망의 일종인 순환 신경망을 사용하였다. 기존의 인공신경망은 입력계층, 은닉계층, 출력계층의 세단계로 구성되며 뉴런들이 단순히 층층이 조합된 구조이다. 이러한 간단한 인공신경망으로도 다양한 분석을 할 수 있으나, 과거의 데이터를 활용해 현재의 값을 예측하는 것에는 한계가 있다. 반면에 순환 신경망은 노드에 과거 정보가 저장되어 다음 시간 단계로 전달되며, 이를 통해 과거 데이터가 현재 값 예측 과정에 사용될 수 있어 시계열 데이터의 분석에 잘 활용된다. 그러나 기본적인 순환신경망은 현재 값 예측에 사용하는 과거 데이터의 길이가 긴 경우 과거의 정보 중 일부가 전달 과정 중에 소실되며, 기울기 소실 문제로 학습이 원활하게 이루어지지 않는다(Gers et al., 2002). 이를 해결하기 위해 LSTM 순환 신경망 알고리즘이 도입되었으며, LSTM은 기존 순환 신경망의 기울기 소실 문제를 LSTM 구조 내에 있는 게이트(gate)들을 사용해 해결한다(Perez-Ortiz et al., 2003).

2.3 LSTM 순환 신경망 알고리즘

LSTM 순환 신경망은 기존의 순환 신경망에서 은닉 상태를 계산하기 전에 LSTM 알고리즘을 거치는 순환 신경망이다. LSTM의 구조로는 여러 모델이 제안되었고, 이 연구에 사용한 모델이 Fig. 1에 나타나있다. 그림에서 x_t, h_t 는 각각 시간 t 에서의 입력과 은닉상태를 나타낸다. i, f, o, g 는 각각 입력 게이트(input gate), 망각 게이트(forget gate), 출력 게이트(output gate), 입력 변조 게이트(input modulation gate)를 나타낸다. 이 연구에서는 Fig. 1에 보이는 한 시간 단계에서의 LSTM 노드가 좌우로 연속적으로 연결되어 있으며, 각 시간 단계에서는 전 시간 단계에서의 은닉상태와 셀 상태를 전달받고, 현 시간 단계에서의 입력값을 받아 게이트들을 통한 연산을 거쳐 은닉상태와 셀 상태를 업데이트 하여 다음 시간 단계로 전송하는 모델을 사용하였다. 가중치(weight)와 편향(bias) 텐서를 각각 W 와 b 로 나타내었다. 사용된 활성화함수인 시그모이드(sigmoid) 함수와 소프트사인(softsign) 함수는 식 (3), (4)로 정의된다. 시그모이드 함수를 거치면 값이 0~1 사이로, 소프트사인 함수를 거치면 값이 -1~1 사이로 변환된다. LSTM 셀 내에서의 계산 과정은 식 (5)~(10)에 나타나있다.

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

$$softsign(z) = \frac{z}{1+|z|} \tag{4}$$

$$i_t = sigmoid(W_{xi}x_t + W_{hi}h_{t-1} + b_i) \tag{5}$$

$$f_t = sigmoid(W_{xf}x_t + W_{hf}h_{t-1} + b_f) \tag{6}$$

$$o_t = sigmoid(W_{xo}x_t + W_{ho}h_{t-1} + b_o) \tag{7}$$

$$g_t = softsign(W_{xg}x_t + W_{hg}h_{t-1} + b_g) \tag{8}$$

$$c_t = f \times c_{t-1} + i_t \times g_t \tag{9}$$

$$h_t = o_t \times softsign(c_t) \tag{10}$$

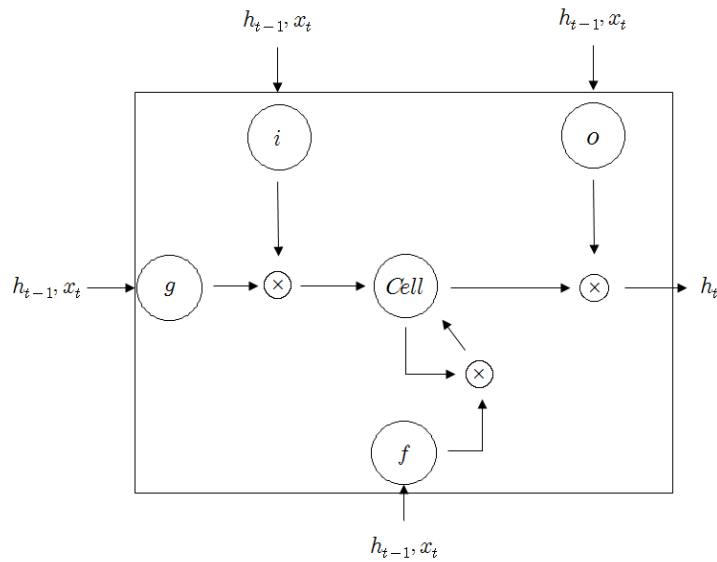


Fig. 1. Internal structure of LSTM block (Zaremba et al., 2014)

본 연구에서 x_t 는 응력 단계 t 에서의 변형률 값을 사용했다. 즉, 신경망에 사용된 입력 자료는 일련의 변형률 값들이다. 이 변형률 값들의 길이인 sequence length를 N 이라 할 경우 입력 자료는 $[x_1, x_2, \dots, x_N]$ 이 되며, 식 (5)~(10)에 따라 $[h_1, h_2, \dots, h_N]$ 이 구해지는데, 그 과정은 Fig. 1과 같이 x_t 와 h_{t-1} 이 입력되어 h_t 가 얻어지는 과정이 순차적으로 $t = 1$ 에서 $t = N$ 까지 반복되는 형태이다(단, $h_0 = 0$). 본 연구에서는 마지막 h 값인 h_N 을 식 (11)의 최종 연산을 통과시켜 나온 값으로 다음 단계에서의 변형률을

예측하였다. 즉, 신경망의 출력(output)인 예측된 변형률 y 는 x_{N+1} 의 추정값이며 식 (11)로 구한다.

$$y = W_{hy}h_N + b_y \quad (11)$$

2.4 데이터 전처리와 모델 학습 과정

LSTM 순환 신경망을 이용해 재료에 특정 응력이 가해졌을 때의 변형률을 예측하였다. 신경망에 입력된 입력 자료는 출력 시점 직전의 일련의 변형률 값들이다. 공업용 반수석고와 물을 섞어 만들어진 석고에 규사를 혼합해서 만든 인공 재료 53개에 대해 0.4 mm/min의 조건으로 하중을 가한 일축압축시험의 데이터(Byun, et al., 2018)를 신경망 모델 구성에 활용했다. 이 시료들은 석고에 세 가지 입도범위(Coarse: 0.76-1.08 mm; Medium: 0.49-0.76 mm; Fine: 0.20-0.49 mm)의 규사 알갱이들을 10%~50%의 비율 범위에서 10%단위로 혼합해 만든 재료들이다. 종합하면 이 재료들은 규사의 크기 3종류와 혼합비율 5종류를 곱한 15가지, 그리고 순수한 석고 1종류를 합한 총 16종류의 재료로 볼 수 있다. 이들의 역학적 특성들의 범위가 Table 1에 표기되었다.

이 연구에 사용하기 위해 일축압축시험을 통해 얻은 초기 데이터(raw data)는 0.2초 단위로 응력과 변형률을 측정된 데이터이다. 이 초기 데이터를 신경망에 입력 자료로 사용하기 위한 전처리 단계는 다음과 같다. 우선 측정기기의 응력과 변형률 측정 시 발생하는 기계의 오차를 줄이기 위해 0.2초 간격을 갖는 연속된 두 개의 수치를 평균하여 사용하였다. 그 후, 응력을 0 MPa부터 파괴시점의 응력까지의 범위에 대하여 0.08 MPa 단위로 나누고 각 응력수준에 따른 변형률 값을 추출해냈다. 한 구간의 길이인 0.08 MPa는 전처리 전후에 데이터셋이 담고 있는 정보의 양과 신뢰성을 비슷하게 유지하기 위한 목적으로 전처리 전후의 데이터셋의 길이를 비슷하게 만들기 위해 선택되었다. 실험으로 측정된 응력들이 0.08 MPa의 정수배가 아니기 때문에 특정 응력에 대응하는 변형률 값의 실험 데이터가 없는 경우가 있다. 이 경우에는 Cubic Spline 보간법을 사용하여 데이터를 생성해냈다. 마지막으로 입력 자료로 사용할 변형률 값에 100을 곱하여 변형률이 0에서 1사이의 분포를 갖게 만들어 학습의 효율을 높였다. 이러한 방식으로 총 53개의 데이터셋에 대해 응력이 0.08 MPa씩 증가할 때 그에 따른 변형률 값들을 얻었다. Fig. 2는 전처리를 마친 53개 데이터셋 중 하나를 그래프로 나타낸 예이다.

Table 1. Mechanical properties of the materials used in this study (Byun, et al., 2018)

UCS (MPa)	Young's modulus (GPa)	Density (g/cm ³)	P-wave velocity (m/s)	S-wave velocity (m/s)
22.7 - 44.0	7.12 - 11.5	1.75 - 2.11	3,269 - 3,922	1,800 - 2,170

실험에 사용된 재료들의 일축압축강도는 25~40 MPa에 속하기 때문에 한 시료의 응력-변형률 데이터셋에 속한 데이터의 갯수(길이)는 이를 0.08 MPa로 나눈 값인 300~500 정도이다. 이 연구에서는 응력들 간의 등간격을 time step으로 설정하고 앞 구간의 변형률을 입력 자료로 하여 뒤쪽 구간의 변형률을 예측하였다. 입력의 크기(길이)인 sequence length는 모델 성능의 비교를 위해 30, 60, 90의 세 종류를 사용하였다. sequence length는 다음 단계의 변형률을 예측하기 위해 사용하는 직전 단계 변형률의 개수를 의미한다. 예를 들어 sequence length값이 30인 경우의 모델의 예측 절차는 다음 순서로 나타낼 수 있다. 이 과정은 시료의 파괴응력까지 이어진다.

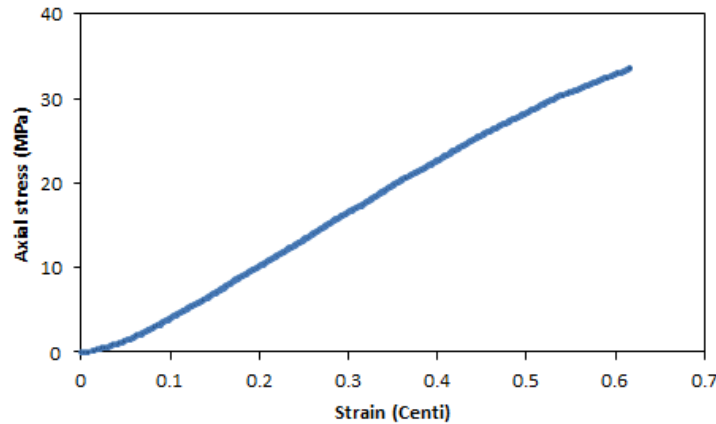


Fig. 2. Preprocessed stress-strain data of the specimen containing 20% coarse grains

1. 0 MPa부터 2.32 MPa까지 구간에서의 변형률을 이용해 2.4 MPa에서의 변형률을 예측
2. 0.08 MPa부터 2.4 MPa까지 구간에서의 변형률을 이용해 2.48 MPa에서의 변형률을 예측
3. 0.16 MPa부터 2.48 MPa까지 구간에서의 변형률을 이용해 2.56 MPa에서의 변형률을 예측

실제 응력-변형률 곡선을 예측할 때는 응력 2.32 MPa까지의 변형률은 일축압축시험을 통해 얻은 실측값을 사용했으며 이후의 변형률을 입력 자료로 사용할 때에는 전 단계에서 예측된 값을 사용하였다. 같은 방법으로 sequence length가 60, 90인 경우도 실제 일축압축시험을 통해 얻은 변형률을 입력 자료로 사용한 구간은 각각 0 MPa부터 4.72 MPa까지와 0 MPa부터 7.12 MPa까지이며 이후의 변형률은 전 단계에서 예측된 값을 사용하였다.

모델 구축과 실험 수행은 Python과 Tensorflow 패키지를 사용하였다. 하이퍼파라미터는 학습 시간과 정확도를 고려하여 시행착오를 거쳐 설정하였다. LSTM 셀에서의 망각 게이트의 편향은 0.4를 사용했으며, hidden unit의 개수는 효율성을 고려해 20으로 하였다. 53개의 실험 데이터는 15종류의 석고+규사 조합의 재료 3개씩(한 종류는 4개)과 순수 석고 시료 7개로 구성되어있다. 따라서 테스트 셋의 대표성을 위해 16종류의 재료를 모두 포함하는 전체 데이터의 약 1/3인 17개의 테스트 셋을 구성하였다. 약 2/3인 나머지 36개의 데이터는 학습에 사용하였다. 최적화 함수로는 Adam Optimizer를 사용하였으며, 드롭아웃(dropout) 비율은 0.01로 설정한 후, 0.0001의 학습률로 300번 이상의 반복 학습을 하였다.

3. 결과 및 토의

3.1 점탄성 모델을 활용한 응력-변형률 곡선 예측

우선 점탄성 모형을 사용하여 응력-변형률 곡선을 예측해 보았다. 53개의 데이터셋 중 하나의 시료에 대해 Burger 모형을 사용한 최적화를 두 가지 경우에 대해 테스트하였다. 두 경우는 각각 전구간 응력-변형률 데이터를 사용한 최적화와, LSTM 순환 신경망을 사용한 시험과 같이 초기 2.32 MPa까지의 데이터만을 사용한 최적화이다.

식 (12)은 전 구간 데이터를 사용한 최적화 결과이며, 이를 그래프로 나타내면 Fig. 3과 같다. 파란색 곡선은 실측값들이며, 회색

곡선이 Burger 모형을 통해 생성한 곡선이다. 경계조건을 만족시키도록 설정하였기 때문에 예측된 응력은 전체 응력의 범위를 만족한다. 하지만 눈에 띄는 오차가 존재하며, Burger 모형으로 응력-변형률 곡선을 표현하는데 한계가 있음을 알 수 있다.

$$\sigma = 0.003386 - 4.193\exp(-997169\epsilon) + 4.18961\exp(330\epsilon) \tag{12}$$

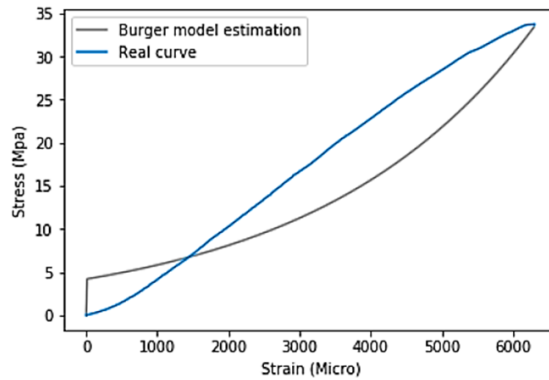
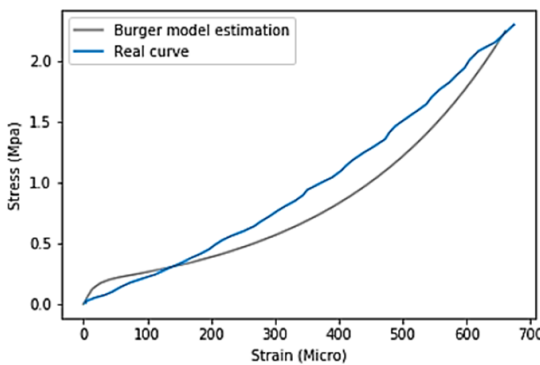


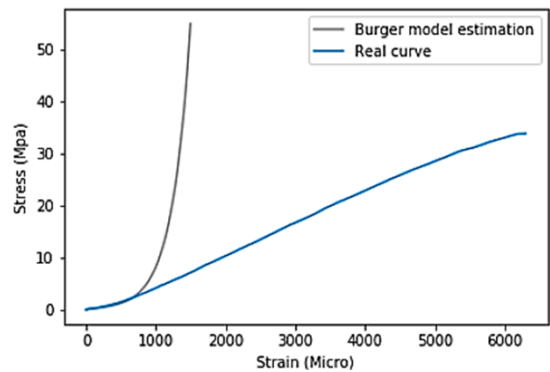
Fig. 3. An example of the presentation of stress-strain curve using Burger model

식 (13)과 Fig. 4는 일부의 데이터셋에 대해 Burger 모형의 변수 최적화를 수행한 결과이다. Fig. 4의 (a)에서 파란색 곡선은 최적화에 사용한 데이터이며, 회색 곡선은 최적화를 통해 설정된 Burger 모형의 변수들을 사용하여 표현한 곡선이다. 데이터의 길이가 크지 않기에 Burger 모형을 사용한 예측이 상당히 정확성을 보이나, 이를 전체 응력 범위에 대해 확장한 Fig. 4의 (b)를 통해, 계산에 사용하지 않은 범위의 응력에서 예측된 변형률은 실제 변형률과 매우 차이가 남을 알 수 있다. 이 과정을 통해 초기의 일부분의 데이터만을 사용해서 모델 변수 최적화를 하는 경우, 향후 곡선 예측에서 큰 불안정성을 보이는 것을 확인하였다.

$$\sigma = 0.000070 - 0.180038\exp(-72451.7\epsilon) + 0.179969\exp(3821.78\epsilon) \tag{13}$$



(a) Initial part of stress-strain curve used for Burger model optimization



(b) Predicted stress-strain curve of the total range of experimental data

Fig. 4. One example of the prediction of stress-strain curve using Burger model with few initial data

3.2 LSTM 순환 신경망을 활용한 응력-변형률 곡선 예측

테스트에 사용한 17개 시료 중 일부의 응력-변형률 곡선의 실측값과 예측값이 Fig. 5~7에 나타나있다. Fig. 5~7의 (a)~(f)는 규사의 크기 세 종류에 대해 함량을 10% 또는 50%로 하는 규사-석고 혼합 재료의 응력-변형률 곡선이다. Fig. 5~7은 sequence length가 각각 30, 60, 90으로 예측에 사용한 초기 단계 데이터의 길이만 다르며, 같은 알파벳의 그래프는 같은 시료로부터 얻어졌다. 가로축은 변형률을, 세로축은 응력을 나타내며, 가로축에 나타난 변형률 값이 모델이 예측한 값이다. Fig. 5~7의 초기 입력 데

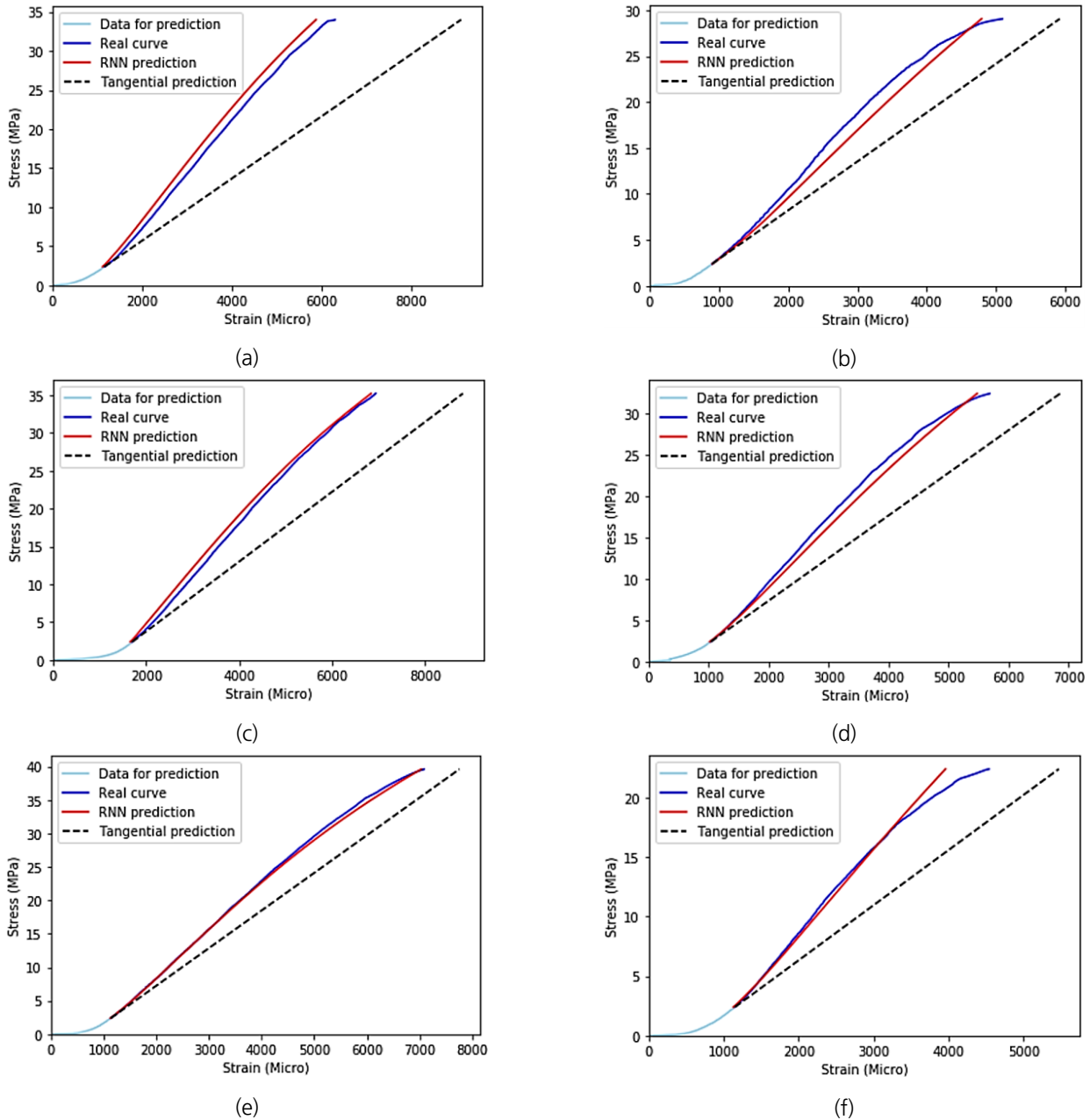


Fig. 5. Predictions of stress-strain curve of the specimens containing (a) 10% coarse grains, (b) 50% coarse grains, (c) 10% medium grains, (d) 50% medium grains, (e) 10% fine grains, and (f) 50% fine grains. Stress range of the data used for prediction is from 0 to 2.32 MPa

이터는 각각 0~2.32 MPa, 0~4.72 MPa, 0~7.12 MPa 구간의 변형률이며 그림에 하늘색으로 표기되어있다. 모델을 통해 예측한 곡선은 빨간색으로 표시되어있으며, 일축압축시험을 통해 얻은 실제 곡선은 파란색으로 나타나있다. 직관적인 예측 방법인 앞쪽 구간의 곡선에 접선을 그리는 방법을 통해 얻어진 응력-변형률 곡선도 이 모델의 실효성을 확인하기 위해 검은색의 파선으로 함께 나타내었다.

대부분의 경우에 대해 응력이 낮은 구간에서 응력이 높은 구간 보다 변형률이 더 잘 예측되었다(Fig. 5~7). 이는 예측 모델을 만들 때 처음의 응력 구간에서만 실제 변형률을 입력 자료로 사용했기 때문으로 판단된다. 이후 구간에서의 입력 자료로 쓰인 변형률

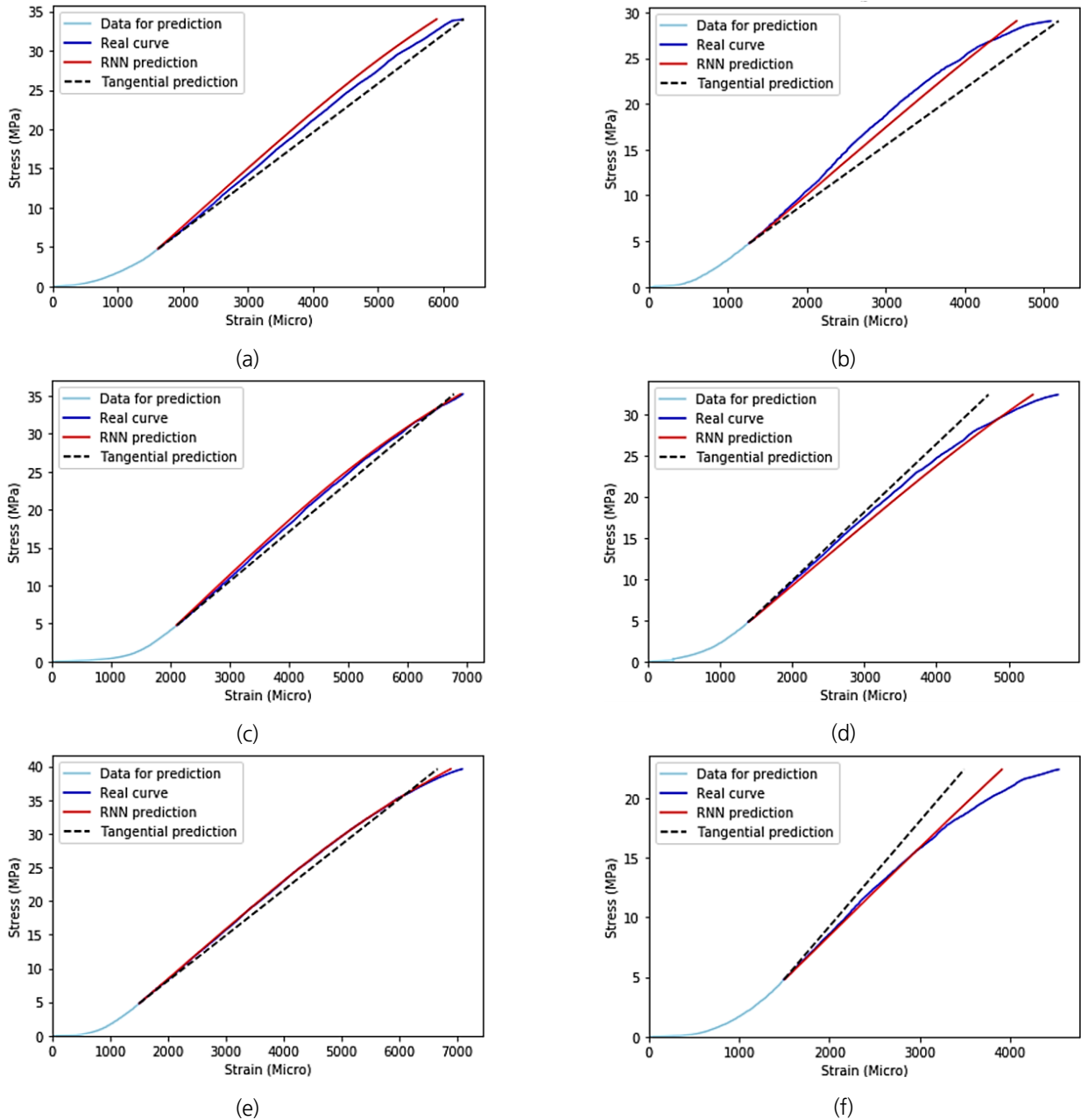


Fig. 6. Predictions of stress-strain curve of the specimens containing (a) 10% coarse grains, (b) 50% coarse grains, (c) 10% medium grains, (d) 50% medium grains, (e) 10% fine grains, and (f) 50% fine grains. Stress range of the data used for prediction is from 0 to 4.72 MPa

은 그 전 단계에서 출력으로 얻어진 변형률이기 때문에 이미 오차를 내포한다. 이 값을 입력 자료로 사용하게 되면 오차가 점점 늘어나며, 응력이 높은 구간일수록 이 오차가 누적되어 실측값과 더 큰 차이를 낸다. 그 결과 높은 응력에서의 비선형 구간은 예측의 정확도가 상대적으로 떨어졌다.

모델의 성능을 정량적으로 판단하기 위해 테스트 셋에 대한 예측의 정확도를 평균제곱근오차(Root Mean Square Error, RMSE)와 결정계수(coefficient of determination)로 나타내었다. 결정계수는 변형률의 예측값과 실측값을 각각 좌표평면에 x, y

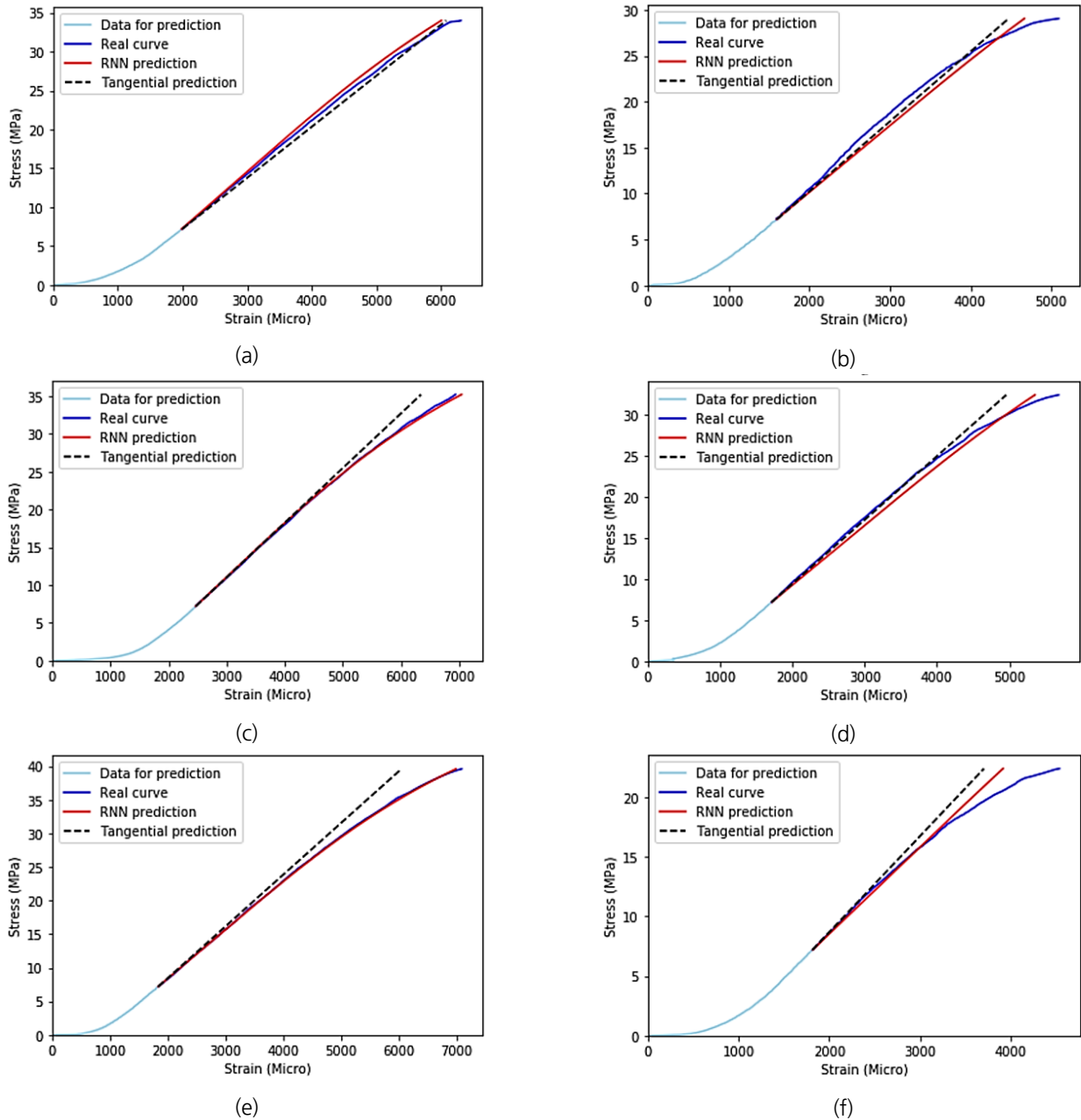


Fig. 7. Predictions of stress-strain curve of the specimens containing (a) 10% coarse grains, (b) 50% coarse grains, (c) 10% medium grains, (d) 50% medium grains, (e) 10% fine grains, and (f) 50% fine grains. Stress range of the data used for prediction is from 0 to 7.12 MPa

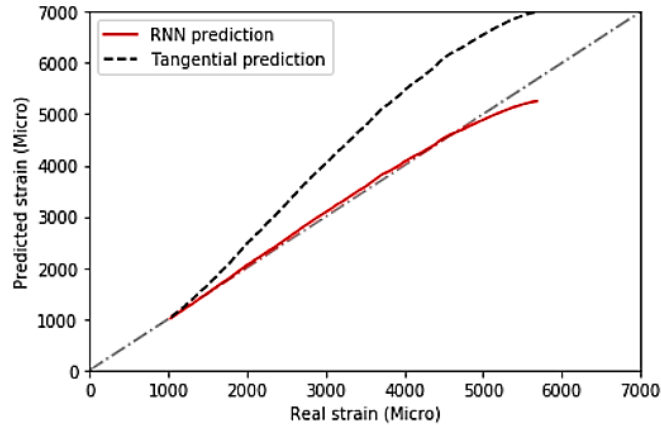


Fig. 8. Comparison between real and predicted strain data of the specimen containing 50% medium grains

값으로 놓고 직선 $y = x$ 와의 일치도로 계산하였다(Fig. 8). 전체 17개의 테스트 셋에 대해 평균제곱근오차와 결정계수를 계산한 결과를 sequence length에 따라 각각 Table 2~4에 나타내었다. 대부분의 경우에 신경망을 사용한 예측 결과가 접선을 활용한 예측 결과보다 정확도가 높게 측정되었으며, 신경망을 사용한 예측 곡선은 실제 곡선과 높은 일치도를 보였다.

Table 2~4에 있는 값들 중 17개 시료에 대한 평균값만을 Fig. 8에 나타내었다. 테스트 셋 17개의 최대 응력의 평균값은 33.51 MPa이므로 초기 입력 데이터의 상한인 2.32 MPa, 4.72 MPa, 7.12 MPa는 전체 데이터에서 차지하는 비율이 각각 6.92%, 14.1%, 21.2%이다. 즉, 각각의 시험은 시료 강도의 약 6.92%, 14.1%, 21.2%까지의 변형률 값으로부터 파괴응력까지의 변형률 값을 예측

Table 2. Performance of the RNN model compared to tangential prediction, both of whose initial input data range is from 0 to 2.32 MPa

Specimen No.	RNN prediction		Tangential prediction	
	R^2	RMSE (Micro strain)	R^2	RMSE (Micro strain)
1	0.9772	211.7	0.3764	1786
2	0.9925	122.6	0.5694	1229
3	0.9834	170.9	0.6241	1021
4	0.9679	231.6	0.9043	413.4
5	0.9722	179.8	0.6247	823.7
6	0.9916	132.8	0.5734	1246
7	0.9804	192.6	0.4809	1415
8	0.9960	90.01	0.7423	839.7
9	0.9902	131.3	0.9714	226.8
10	0.9874	138.4	0.6035	992.9
11	0.9984	65.15	0.8387	706.0
12	0.9930	107.0	0.5398	1175
13	0.9672	257.6	0.8056	687.3
14	0.8564	487.1	0.8194	558.6
15	0.9797	127.8	0.5388	821.2
16	0.8960	536.6	0.6213	1230
17	0.9875	156.9	0.9569	296.0
Average	0.9716	196.5	0.6818	909.8

Table 3. Performance of the RNN model compared to tangential prediction, both of whose initial input data range is from 0 to 4.72 MPa

Specimen No.	RNN prediction		Tangential prediction	
	R^2	RMSE (Micro strain)	R^2	RMSE (Micro strain)
1	0.9852	157.7	0.9757	203.0
2	0.9937	105.4	0.9438	322.1
3	0.9754	194.3	0.9402	308.7
4	0.9809	166.1	0.8182	560.9
5	0.9817	133.5	0.8410	425.8
6	0.9982	57.41	0.9912	127.2
7	0.9874	142.5	0.9791	184.3
8	0.9889	141.3	0.9659	250.8
9	0.9798	178.1	0.7668	683.9
10	0.9915	105.7	0.9460	272.8
11	0.9996	30.72	0.9910	144.9
12	0.9977	55.77	0.8823	427.2
13	0.9683	238.0	0.9754	209.2
14	0.9140	341.6	0.9006	370.0
15	0.9662	150.8	0.8369	355.7
16	0.9195	431.1	0.9573	307.6
17	0.9936	104.9	0.9802	185.4
Average	0.9777	160.9	0.9230	314.1

Table 4. Performance of the RNN model compared to tangential prediction, both of whose initial input data range is from 0 to 7.12 MPa

Specimen No.	RNN prediction		Tangential prediction	
	R^2	RMSE (Micro strain)	R^2	RMSE (Micro strain)
1	0.9933	97.58	0.9942	90.77
2	0.9967	70.85	0.9025	404.6
3	0.9785	168.8	0.9564	243.1
4	0.9804	158.1	0.9502	255.8
5	0.9756	142.9	0.9784	134.4
6	0.9992	36.17	0.9762	196.2
7	0.9940	91.16	0.9822	157.7
8	0.9898	126.8	0.9649	238.5
9	0.9779	174.9	0.8751	439.0
10	0.9887	113.5	0.9748	171.0
11	0.9997	23.76	0.9500	328.3
12	0.9934	87.72	0.8255	493.9
13	0.9561	265.2	0.9315	335.6
14	0.9334	278.3	0.9557	224.4
15	0.9546	158.6	0.8889	257.15
16	0.9523	300.9	0.9609	271.3
17	0.9881	133.9	0.9098	384.1
Average	0.9795	142.9	0.9398	272.1

한 것이다. Fig. 9의 가로축은 이 비율을 나타내며, 세로축은 각 모델의 성능을 나타낸다. 예측에 사용하는 초기 데이터의 길이가 작은 경우(6.92%)는 응력-변형률 곡선 초반의 비선형 구간을 통해 뒤쪽 구간을 예측하는 것이지만, 초기 데이터의 길이가 긴 경우(21.2%)에는 사실상 초기 구간이 선형 구간에 진입했다고 볼 수 있다. 따라서 예측에 사용하는 초기 데이터의 길이를 늘리면 이 부분이 선형 구간에 진입함에 따라 접선을 통한 예측의 정확도가 급격히 증가하며 신경망 모델을 통한 예측의 정확도와와의 차이가 줄어든다(Fig. 9). 따라서 이 연구에서 실험한 세 경우 모두 신경망을 통한 예측의 정확도는 접선을 통한 예측의 정확도를 상회하나, 초기 데이터의 길이가 짧은 경우에 신경망 모델이 더욱 유용함을 알 수 있다. 결국 이 모델은 응력-변형률 데이터가 매우 적지만 높은 응력에서의 거동을 예측하고 싶은 경우 효과적이다.

이 실험 결과는 재료의 초기단계의 응력-변형률 데이터만을 사용해 파괴까지의 응력에 따른 변형률을 높은 정확도로 예측할 수 있음을 보여준다. 이는 이론적인 계산이나, 전과정 일축압축시험 없이 응력-변형률 곡선을 얻을 수 있다는 점에서 의의가 있다. 이 연구는 재료에 초기상태의 약한 응력이 가해졌을 때 변형을 측정할 수 있을 경우 향후 더 큰 응력이 가해질 상황에서 변형이 어느 정도 일어날지 예측하기 위한 목적으로 활용될 수 있을 것이다.

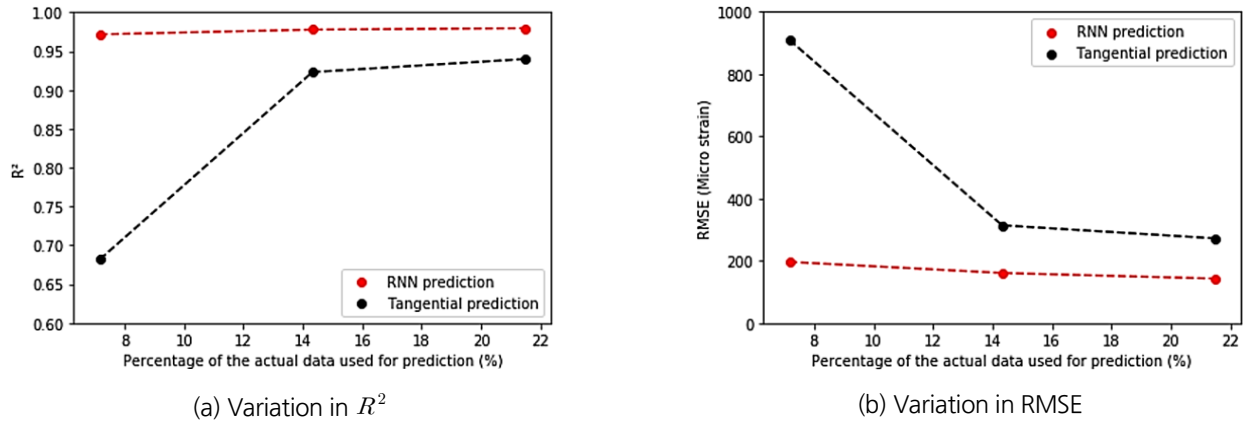


Fig. 9. Performance of RNN and tangential prediction for different percentage of the input data used for prediction

4. 결론

석고와 규사를 혼합한 재료로부터 얻은 일축압축시험 데이터를 통해 응력-변형률 곡선을 예측하는 모델을 수립하였다. 기존에는 재료의 비선형 거동을 해석하기 위해 주로 유변학적 모델을 사용하였으나, 본 연구를 통해 신경망 구조상 시계열 데이터를 다루는데 유리한 순환 신경망을 사용한다면 특정한 모델을 도입할 필요 없이 오직 경험적인 방법만으로 재료의 거동을 묘사할 수 있는 것을 확인하였다. 뿐만 아니라, 일부의 데이터로 전체 응력-변형률 곡선을 예측하는 경우에 높은 불안정성을 띠는 점탄성 모형과 달리, 순환 신경망 모델 도입을 통해 안정적으로 재료의 거동 예측이 가능함을 확인하였다. 순환 신경망 모델 시험에서 대부분의 경우 응력-변형률 곡선은 높은 응력 상태에서는 오차가 조금 생겼으나 낮은 응력 상태에서는 실측값과 비슷한 곡선을 생성해 낼 수 있었다. 또한, 예측에 사용한 초기 데이터의 길이를 증가시킬수록 모델의 정확도는 조금 올라갔으나, 접선 예측방법과의 비교를 통해 예측의 정확성을 상대적으로 비교해본 결과, 적은 초기 데이터를 사용할 경우에 신경망 모델의 특징점이 더욱 부각되는 것을 알 수 있다. 신경망 모델은 매우 적은 데이터를 기반으로 하여도 향후 거동 예측이 가능하다. 재료 강도의 약 7% 이하인 비선형 구간에서의

곡선만으로 이후 응력-변형률 곡선을 예측할 수 있기 때문에, 이 모델은 비선형 거동을 보이는 물체의 초기 응력에 대한 변형률을 알 경우 더 큰 응력 하에서의 변형률을 계산해야하는 상황에 적용될 수 있다.

이 시험은 석고와 규사를 혼합한 재료에 대해서만 수행되었으나, 일반적으로 점탄성을 보이는 많은 재료들의 응력-변형률 곡선은 비슷한 형태를 띤다. 이 논문에서 사용한 모델이 규사 입자의 크기와 함량을 달리하는 다양한 물성을 가지는 재료의 응력-변형률 곡선을 표현하는데 효과적임을 밝혔으므로, 암석 시료와 같이 새로운 물성을 갖는 재료의 거동을 예측하고자 하는 경우에는 이 논문에서 사용한 것과 같은 구조의 신경망 모델을 수립하고 새로운 재료들의 데이터로 학습을 시키면 된다. 또는 이 논문에서 학습이 완료된 모델에 추가로 실제 암석의 실험 결과를 입력 자료로 하여 학습시킨다면, 일반적인 재료의 거동을 예측하는 하나의 모델을 만들 수 있을 것이다. 이러한 것이 가능한 이유는 입력 자료로 사용된 초기 응력-변형률 곡선에 해당 재료의 주요 물성이 반영되어 있기 때문이다.

사사

이 논문은 정부(과학기술정보통신부, 환경부, 산업통상자원부)의 재원으로 한국연구재단-탄소자원화 국가전략프로젝트사업의 지원을 받아 수행함(M3D8A2085654).

REFERENCES

- Arindam, D. and Prabir, B., 2012, Estimation of Burger model parameters using inverse formulation, *International Journal of Geotechnical Engineering*, Vol. 6, No. 3, pp. 261-274.
- Byun, H., Fereshtenejad, S. and Song, J., 2018, A study on the effect of grain content and size on mechanical properties of artificial sedimentary rocks, *Tunnel and Underground Space*, Vol. 28, No. 2, pp. 156-169.
- Gers, F., Schraudolph, N. and Schmidhuber, J., 2002, Learning precise timing with LSTM recurrent networks, *Journal of Machine Learning Research*, Vol. 3, pp. 115-143.
- Guan, Z., Jiang, Y., Tanabashi, Y. and Huang, H., 2008, A new rheological model and its application in mountain tunnelling, *Tunnelling and Underground Space Technology*, Vol. 23, No. 3, pp. 292-299.
- Hardy, H.R., 1967, Analysis of the inelastic deformation of geologic materials in terms of mechanical models, Spring Meeting, Society for Experimental Stress Analysis, Ottawa, Canada.
- Hudson, J.A. and Harrison, J.P., 1997, *Engineering Rock Mechanics: an introduction to the principle*, Pergamon, Elsevier Science, Oxford 215p.
- Karami, M. and Fahimifar, A., 2013, A new time-dependent constitutive model and its application in underground construction, *ISRM International Symposium on Rock Mechanic for Resources, Energy and Environment*, Wroclaw, Poland, pp. 21-26.
- Li, Y., Zhu, W., Fu, J., Guo, Y. and Qi, Y., 2014, A damage rheology model applied to analysis of splitting failure in underground caverns of Jinping I hydropower station, *International Journal of Rock Mechanics and Mining Sciences*, Vol. 71, pp. 224-234.
- Perez-Ortiz, J., Gers, F., Eck, D. and Schmidhuber, J., 2003, Kalman filters improve LSTM network performance in problems

unsolvable by traditional recurrent nets, *Neural Networks*, Vol. 16, No. 2, pp. 241-250.

Skrzypek, J.J., Ganczarski, A.W., 2015, Constitutive Equations for Isotropic and Anisotropic Linear Viscoelastic Materials, *Mechanics of Anisotropic Materials, Engineering Materials*. Springer, Cham, pp 57-85.

Zaremba, W., Sutskever, I. and Vinyals, O., 2014, Recurrent neural network regularization, arXiv preprint arXiv:1409.2329v5.

Zhao, D., Jia, L., Wang, M. and Wang, F., 2016, Displacement prediction of tunnels based on a generalised Kelvin constitutive model and its application in a subsea tunnel, *Tunnelling and Underground Space Technology*, Vol. 54, pp. 29-36.