

Research Article Modelling Hysteresis in Shape Memory Alloys Using LSTM Recurrent Neural Network

Mohammad Reza Zakerzadeh 🕞, Seyedkeivan Naseri 🕞, and Payam Naseri 🕒

School of Mechanical Engineering, College of Engineering, University of Tehran, Tehran, Iran

Correspondence should be addressed to Mohammad Reza Zakerzadeh; zakerzadeh@ut.ac.ir

Received 28 November 2023; Revised 18 January 2024; Accepted 23 January 2024; Published 14 February 2024

Academic Editor: Waqar A. Khan

Copyright © 2024 Mohammad Reza Zakerzadeh et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The complex behavior of shape memory alloys (SMAs), characterized by hysteresis and nonlinear dynamics, results in complex constitutive equations. To circumvent the complexity of solving these equations, a black box neural network (NN) has been employed in this research to model a rotary actuator actuated by an SMA wire. Considering the historical dependence of the pulley's rotational angle on the applied voltage, a recurrent neural network (RNN) is suitable for capturing past information. Specifically, a long short-term memory (LSTM) neural network is selected due to its ability to address issues encountered in standard recurrent networks. There are major drawbacks with modelling hysteresis with NNs that do not account for historical behavior. Traditional NNs, characterized by a one-to-one mapping, struggle to capture hysteresis loops wherein system behavior varies during loading and unloading cycles. Therefore, a single-tag data is used to determine the loading or unloading state, but tag signal causes discontinuity in network and omits various aspects of hysteresis in SMA, particularly within minor loops. In contrast, NNs incorporating past data to predict hysteresis behavior alleviate the need for tag data. However, such networks tend to have complex structures with a substantial number of neurons to effectively capture the inherent nonlinearity in SMAs. The long short-term memory (LSTM) neural network employed in this research, characterized by a simpler structure, achieves high accuracy in predicting hysteresis in SMAs without the need for tag data. In the proposed LSTM model, data related to the pulley's rotational angle and the wire's applied voltage from the current moment and the two previous moments serve as input. The data passes through a layer comprising three LSTM cells, and the output from the last LSTM cell is fed into a fully connected layer to predict the pulley's rotational angle for the next moment. Training data are obtained by applying voltage at various frequencies and formats to the SMA wire while simultaneously recording the pulley's angle with an encoder. Evaluation of the LSTM model is conducted in two configurations: online prediction (one-step ahead) and offline prediction (multistep ahead). In the online configuration where the model uses encoder data as angular inputs, the root mean square error (RMSE) of predictions for various input voltages is significantly low at about 0.1 degrees where the maximum rotational angle of pulley is 8 degrees. In the offline configuration when using the model's predictions as angular inputs instead of encoder data, the RMSE rises to 0.3 degrees. To provide a clear demonstration of the LSTM model's ability in this particular configuration, a comparison has been conducted between LSTM model and a rate-dependent Prandtl-Ishlinskii (RDPI) hysteresis model for predicting the pulley's angle. The LSTM model outperforms the RDPI model by 70% in terms of accuracy. Overall, the LSTM model demonstrates capability in effectively modeling SMA hysteresis in both online and offline configurations.

1. Introduction

Shape memory alloys (SMAs) are a class of materials that exhibit the unique shape memory effect (SME) due to their crystalline structure. SME allows the alloy to recover its strain when its temperature is changed, which induces a phase change between the austenite and martensite crystal structures [1]. This shape memory property along with high force-to-mass ratio, biocompatibility, and silent operation makes SMA ideal for various applications requiring significant force and movement [2, 3].

SMAs exhibit nonlinear dynamics coupled with hysteresis behavior, resulting in complex material characteristics. The complexity is further heightened by the dependence of this behavior on factors such as applied stress, transformation temperature, the percentage of martensite and austenite at any given moment, and constituent elements of the alloy making modeling such alloys even more challenging. Researchers have introduced various modeling approaches to predict SMA behavior, including constitutive models, hysteresis models, and those trained by machine learning (ML) methods.

Constitutive models of SMAs attempt to describe the behavior of these alloys as a function of variables including stress, strain, temperature, and time rate of them. The Tanaka model is one of the earliest constitutive models proposed for SMAs, introduced in 1986 [4]. In this model, strain, temperature, and the volume fraction of martensite phase are considered as state variables, and stress is calculated as a function of these variables. Additionally, the phase transformation kinetics is expressed exponentially and is a function of stress and temperature. Liang and Rogers [5] built upon Tanaka's research to introduce a novel set of empirical equations for phase transformation kinematics. Their approach involves a simplified kinetic relation, represented by a cosine function, to describe the martensitic phase fraction. While the Tanaka and Liang-Rogers models successfully describe the phase transformations between martensite and austenite, they are limited in that they do not account for the detwinning of martensite that produces the SME at lower temperatures [6]. The Brinson model [7] was utilized to solve this problem. In the Brinson model, the martensite volume fraction is separated into stressinduced and temperature-induced components. There are also modified versions of the Brinson model [8], but the complexity of the equations increases along with the model's accuracy in predicting SMA behavior.

Hysteresis in SMAs refers to the phenomenon where the material exhibits different behaviors during loading and unloading cycles so the material's response is history-dependent, meaning it depends not only on the current input but also on its previous states. Hysteresis models can be classified into two main categories: operator-based models like Preisach [9], Krasnoselskii-Pokrovskii [10], and Prandtl-Ishlinskii [11] that use play operators and differential equation-based models [12, 13]. Operator-based models can accurately predict hysteresis but require complex computations. Differential equation-based models are simpler but less flexible in modeling complex hysteresis.

Constitutive models for SMAs use complex equations to describe hysteresis and nonlinear stress-strain-temperature relationships. Determining these equations is timeconsuming. On the other hand, hysteresis models only consider one input parameter, neglecting others that constitutive models include. To overcome these challenges, ML methods have been proposed as an alternative to model SMA behavior.



FIGURE 1: SMA and rotary encoder connected to the pulley.

In recent years, ML methods have been applied to various applications in real life. In health care field, ML algorithms have been instrumental in diagnosing diseases and predicting patient outcomes with greater accuracy [14, 15]. In the finance industry, ML has enabled professionals to predict financial parameters with precision, leading to better investment decisions and risk management [16]; in the field of geology, ML enables professionals to analyze vast amounts of geological and spatial data to make informed decisions on environmental planning [17–19] and urban and rural development [20].

Neural networks (NNs), as a subset of ML, have proven effective in representing the hysteresis characteristics. NNs are a viable alternative to traditional modeling approaches for capturing the complex behavior of SMAs. In a 2003 study [21], researchers used a shallow NN as an open-loop controller to tracking control of an SMA actuator. The inputs to the NN were desired outputs and a label indicating whether the system was in the heating or cooling state. In another 2010 study [22], researchers used a shallow NN to estimate the strain of an SMA wire. The inputs to this NN were the resistance of the wire at each moment and binary values indicating whether the system was in the heating or cooling state, but this approach requires the SMA to be only on major hysteresis loops, meaning SMA should be fully expanded or fully contracted. This NN was then used in a proportional-integral-derivative (PID) control algorithm to estimate the displacement and consequently eliminate the need for a displacement sensor. In 2011, Zakerzadeh et al. used an NN to approximate functions that determined the hysteretic behavior of a numerical Preisach model. The results demonstrated that NNs for numerical function approximation provide higher accuracy in predicting hysteresis behavior compared to the classical Preisach model and numerical approaches [23]. In 2013, Wang and Song introduced a new type of recurrent neural network (RNN) that can predict the hysteresis behavior of an SMA wire at different frequencies. The output of this NN was the strain of the wire in the next moment, and its inputs included the previous output values of the NN and the given current value of electrical current [24]. In a 2018 study [25], researchers



FIGURE 2: Data obtained from experimental setup. (a) Input signal $v_1(t)$ with f = 0.03 Hz and $\tau = 0.008$ s. (b) Input signal $v_2(t)$ with f = 0.03 Hz and $\tau = 0.008$ s. (c) Response of SMA-actuated system to $v_1(t)$ with f = 0.03 Hz and $\tau = 0.008$ s. (d) Response of SMA-actuated system to $v_2(t)$ with f = 0.03 Hz and $\tau = 0.008$ s. (e) Response of SMA-actuated system to $v_1(t)$ with f = 0.03 Hz and $\tau = 0.008$ s versus duty cycle. (f) Response of SMA-actuated system to $v_2(t)$ with f = 0.03 Hz and $\tau = 0.008$ s versus duty cycle.

predicted the displacement of an SMA spring using an NN. The NN used in this study was a feedforward network with 3 hidden layers and 11 neurons in the input layer. The inputs to the NN included the electrical current, force, and temperature at the current moment and two previous moments, as well as the amplitude and frequency of the SMA electrical current. In 2020 [26], researchers used two different NNs to predict the temperature of an SMA wire. The first NN was a feedforward network with two hidden layers with 32 and 16 neurons, and its inputs were the differential resistance value, four current values, and a label determining whether the input voltage was increasing or decreasing. The second NN used for predicting the temperature of the SMA wire was a long short-term memory (LSTM) network, and its inputs were the current values and the differential resistance values up to three previous moments. The second NN achieved significantly higher accuracy. In 2022, researchers used an innovative NN to estimate the displacement of an SMA actuator consisting of a pair of antagonistic SMA wires [27]. The NN used in this research consisted of three parts. In the first part, an LSTM neural network was used, with the input being the differential resistance values of the wire in the last 50 moments and the output being the temperature values of the wire in the last 50 moments. In the second part, a feedforward network was used to model the static relationship between the temperature value and the martensitic volume fraction of the SMA wire. In the third part, similar to the first part, an LSTM network was used, with the inputs being the martensitic volume fraction values in the last 50 moments and the output being the displacement value at the current moment. The results obtained in this study were compared with the results of a 2-layer LSTM neural network, and it demonstrated that the designed network in this research provides better results.

The aim of this research is to construct a model for a rotary actuator actuated by SMA using LSTM neural networks. In contrast to previous works that utilized NNs, using LSTM network eliminates the need for single-tag data to determine whether the SMA wire is in a loading or unloading state. Furthermore, LSTM networks demonstrate the capability to model both major and minor hysteresis loops [21, 22]. Additionally, owing to time series nature of LSTM network, there is no requirement for supplemental information such as the frequency of the input signal [25]. Using LSTM network allows to use simpler architecture thereby circumventing the need for multiple feedback loops to capture historical relations in hysteresis loops [24].

The paper is structured as follows. Section 2 presents the experimental setup of the SMA-actuated rotary actuator and introduces the input signals used to obtain training data for the LSTM model. In Section 3, the proposed LSTM model is introduced. Section 4 presents the performance of the proposed model and compares the results with a rate-dependent Prandtl-Ishlinskii (RDPI) hysteresis model. Finally, in Section 5, we conclude the outcomes and goals of the research.

2. Experimental Setup

The test setup shown in Figure 1 consists of a pulley with a radius of 2 cm and mass of 0.05 kg, actuated by two antagonistic SMA wires. SMA wires are connected directly from the pulley to a fixed base. In each moment, one of SMA wires is heated through voltage applied to its terminals and the



FIGURE 3: RRN structure unrolled in time.

other wire is initially contracted and serves as a spring, generating an opposing moment against the first wire.

The SMA wires used in this research are of the Flexinol type, having a diameter of 0.008 inches and a length of 50 cm. The pulleys can withstand temperatures up to approximately 200 degrees Celsius. It is worth noting that the temperature of the SMA wires will never exceed 160 degrees Celsius to prevent damage. A 3600-pulse rotary encoder (Autonics - E50S Series) measures the rotational angle of the pulley. To apply current to the SMA wires, an Arduino control board, a single-channel power supply with a maximum output of 32 volts and 3 amps, and a motor driver (LMD5560) which regulates and switches the current from the power supply to the SMA wires are used.

The input to the system is a pulse-width modulation (PWM) voltage signal applied to the SMA wire's terminals. The output of the system is the angle of rotation of the pulley sampled at a rate of 20 Hz. To better model the behavior of the SMA wire, we use two types of inputs to obtain the required NN training data. In the first type (Equation (1)), the input value reaches zero in each cycle, while in the second type (Equation (2)), the input value is nonzero in each cycle. The two input types are as follows:

$$v_1(t) = Ae^{-\tau t} \left(\sin \left(2\pi f \ t - \frac{\pi}{2} \right) + 1 \right),$$
 (1)

$$v_2(t) = Ae^{-\tau t} \left(\sin \left(2\pi f \ t - \frac{\pi}{2} \right) \right) + A. \tag{2}$$

The input signals have a sinusoidal form that decrease in amplitude over time with a decay time constant of $\tau = 0.008$. In this research, parameter A in Equation (1) and Equation (2) is set to 6 V, so the max value of signals $v_1(t)$ and $v_2(t)$ does not exceed 12 V. Therefore, the PWM duty cycle ranges 0-100% for 0-12 V. The frequency (*f*) of the input signal varies from 0.03 to 0.07 Hz across trials. Figures 2(a) and 2(b) display examples of input signals for Equation (1) and Equation (2), respectively, both with a frequency of 0.03 Hz. Figures 2(c) and 2(d) then show the corresponding system response to these inputs. Finally, Figures 2(e) and 2(f) illustrate the system response versus duty cycle for Equation (1) and Equation (2) when using the 0.03 Hz input signals.

3. Modelling

Output of a system that exhibits hysteresis behavior depends not only on the current input but also on previous inputs—in









Time window size	1	2	3	5	10	20	50
RMSE (deg)	0.0809	0.0571	0.0380	0.0398	0.0435	0.0421	0.0510

other words, the system has a form of memory. As shown in Figure 3, RNNs have a feedback loop, where the network's output is fed back into the network along with the next input. This allows RNNs to retain information about previous inputs in their internal memory, which can then be used to process sequential inputs. In essence, the feedback loop in RNNs gives them a form of memory derived from persisting previous inputs. 3.1. Recurrent Cell. RNNs are often built using standard recurrent cells, such as sigmoid and tanh units. Figure 4 shows a diagram of a typical recurrent sigmoid unit. The mathematical equation defining this standard recurrent tanh cell is

$$h_t = \tanh(W_h h_{t-1} + W_x x_t + b),$$
 (3)



FIGURE 6: Architecture of the proposed RNN model.

where x_t and h_t are input and recurrent information at time t, respectively, W_h and W_x are weights, and b is the bias. Standard RNNs with conventional recurrent units struggle with long-term dependencies; the larger the gap between relevant inputs, the harder it is to learn connections between them. Analyses by Hochreiter and Schmidhuber in [28] and Bengio et al. in [29] identified key reasons for this long-term dependency problem: error signals propagating backward through time tend to either explode or vanish.

The original LSTM was proposed by Hochreiter and Schmidhuber in 1997 [28] to address the problem of learning long-term dependencies in sequence data that happens in standard recurrent cell. The main difference between LSTM and standard RNN is the structure of the LSTM cell. Figure 5 illustrates an LSTM cell which contains an input gate, output gate, and forget gate. The gates act as pathways to control the flow of information, allowing only relevant data to pass through.

The mathematical expressions that define the LSTM cell, as shown in Figure 5, are as follows:

$$f_t = \sigma \left(W_{fh} h_{t-1} + W_{fx} x_t + b_f \right), \tag{4}$$

$$i_t = \sigma(W_{ih}h_{t-1} + W_{ix}x_t + b_i),$$
 (5)

$$\overline{C}_t = \tanh\left(W_{\overline{c}h}h_{t-1} + W_{\overline{C}x}x_t + b_{\overline{C}}\right),\tag{6}$$

$$C_t = f_t \cdot C_{t-1} + i_t \cdot \bar{C}_t, \tag{7}$$

$$o_t = \sigma(W_{oh}h_{t-1} + W_{ox}x_t + b_o), \tag{8}$$

$$h_t = o_t. \tanh(C_t), \tag{9}$$

where operator "." denotes pointwise multiplication of two vectors. The forget gate (f_t) in particular allows the LSTM to forget or retain information in the cell state. It assigns values between 0 and 1 to each element in the cell state, determining what information to keep or forget. This helps the LSTM model retain long-term dependencies that are useful while forgetting irrelevant past data. The input gate determines whether information should be retained or discarded in the LSTM cell. It consists of two components that



FIGURE 7: Loss of LSTM network during training.

TABLE 2: LSTM network hyperparameters.

Optimization algorithm	Learning rate	Batch size	Number of epochs
Adam	0.002	64	40

are multiplied together: i_t and \overline{C}_t . The previous hidden state and current input pass through a sigmoid activation (Equation (5)) to create i_t , while the previous hidden state and current input go through a tanh activation (Equation (6)) to create \overline{C}_t . The cell state (C_t) is defined by Equation (7), where the previous cell state is multiplied by the forget gate and the result is added to the input gate. The output gate also has two parts. First, the previous hidden state and input pass through a sigmoid activation (Equation (8)) to create o_t . Second, the output is multiplied by the cell state passing through a tanh activation (Equation (9)).







FIGURE 8: Actual and predicted angle by LSTM model in one-step-ahead prediction configuration. (a) I_1 as input signal. (b) I_2 as input signal. (c) I_3 as input signal. (d) I_4 as input signal. (e) I_5 as input signal.

3.2. Proposed Model. In this research, an LSTM neural network has been used to model the hysteresis behavior of the system with an SMA actuator. The input (S_t) to the LSTM neural network contains the duty cycle values of the input voltage signal and the angular pulley rotation for the past sliding window of time. The output of the model is the pulley angle for the next moment. The number of previous data points considered for predicting the pulley angle is determined by the sliding window size. If we consider the sliding time window size as d, meaning for predicting each output, we look at the past d system moments, then the NN input at each moment is a $d \times 2$ matrix which is shown in the following equations:

$$S_t = [X_t, X_{t-1}, \cdots, X_{t-d-1}],$$
 (10)

$$X_t = \begin{bmatrix} v_t \\ \theta_t \end{bmatrix},\tag{11}$$

where X_t is a vector consisting of duty cycle of voltage applied to SMA wire (v_t) and angle of pulley (θ_t) .

To determine the optimal time window size, d, the NN has been trained using different values of d. Prediction errors on evaluation data have been analyzed, as shown in Table 1. The results indicate that as d increased from 1 to 3, the prediction error decreases significantly. However, further increasing d not only stops decreasing the error but slightly increases it. Given these observations, we can conclude that the most suitable time window size is d = 3. The architecture of the proposed model is presented in Figure 6.

At each time step, the model takes the current and previous 2 time steps of data on the pulley's rotational angle as well as duty cycle of the input voltage as input. The input

TABLE 3: Errors of LSTM model in one-step-ahead prediction configuration for five types of inputs.

Input	I_1	I_2	I_3	I_4	I_5
RMSE (deg)	0.0559	0.0668	0.1045	0.1096	0.0490

data passes through 3 LSTM layers, with the output of the last LSTM layer being the hidden state vector h_t . This vector h_t is then fed into a fully connected layer with 64 neurons. The output from this fully connected layer gives the predicted rotational angle for the next time step. The network was trained offline using the neural network toolbox in MATLAB software. The training and validation losses are shown in Figure 7, while the remaining hyperparameters of the network can be found in Table 2.

Once the training of the LSTM model with the data specified in Section 2 is completed, its performance is assessed by evaluating it on 5 sets of test data that were not used for training. The inputs for test sets are generated using equations (12)-(16):

$$I_1(t) = Ae^{-0.008t} \left(\sin \left(2\pi (0.07)t - \frac{\pi}{2} \right) + 1 \right), \tag{12}$$

$$I_2(t) = Ae^{-0.01t} \left(\sin \left(2\pi (0.07)t - \frac{\pi}{2} \right) + 1 \right), \tag{13}$$

$$I_3(t) = Ae^{-0.01t} \left(\sin \left(2\pi (0.07)t - \frac{\pi}{2} \right) \right) + A.$$
(14)

Furthermore, a signal with a variable frequency has been used as follows:

$$I_4(t) = Ae^{-0.01t} \left(\sin \left(2\pi (0.07 - 0.0002t)t - \frac{\pi}{2} \right) + 1 \right).$$
(15)



FIGURE 9: Continued.



FIGURE 9: Actual and predicted angle by LSTM model in multi-step-ahead prediction configuration. (a) I_1 as input signal. (b) I_2 as input signal. (c) I_3 as input signal. (d) I_4 as input signal. (e) I_5 as input signal.

TABLE 4: Errors of LSTM model in multi-step-ahead prediction configuration for five types of inputs.

l'able 5: RDF	'l model	parameter	values
---------------	----------	-----------	--------

Input set	I_1	I_2	I_3	I_4	I_5
RMSE (deg)	0.1807	0.3509	0.2269	0.2889	0.1518

Since an ascending signal has not been considered in the training data, the following equation is used to consider an ascending sinusoidal signal as the test signal:

$$I_5(t) = Ae^{0.008t} \left(\sin \left(2\pi (0.03)t - \frac{\pi}{2} \right) + 1 \right).$$
 (16)

3.3. One-Step-Ahead Prediction. The system model receives inputs I_1 through I_5 . The model's input consists of the pulley's angle and duty cycle at the current and two previous moments. This input is updated at each moment with the actual pulley angle value recorded by the encoder. This allows the model to predict the pulley's angle for the next moment. Figure 8 presents the model's pulley angle predictions and actual pulley angle values for the different inputs. Table 3 shows the model's root mean square error (RMSE) values for one-step-ahead predictions across the five input groups I_1 to I_5 .

3.4. Multi-Step-Ahead Prediction. In this section, we use the model to predict the system's response without having its actual values at each moment. For this purpose, the model's predictions are provided as inputs to the system for subsequent time steps. In this case, a specific measured angle of zero degree is used as the initial value. The model is utilized in this setup to predict all five groups of test data given in equations (12)–(16). The results are presented in Figure 9.

RDPI model parameters	Values
N	10
C_1	13.03
C_2	0.02
C_3	0.14
C_4	-0.99
C ₅	12.72
C ₆	0.01
C_7	-1.11
C_8	1.02
<i>C</i> ₉	13.46
C_{10}	0.17
9	0.07
<i>a</i> ₁	16.57
<i>a</i> ₂	-39.98
<i>a</i> ₃	26.55
a_4	9.87
<i>a</i> ₅	30.19
<i>a</i> ₆	26.16
<i>a</i> ₇	34.2
<i>a</i> ₈	7.85
<i>a</i> ₉	-21.83
<i>a</i> ₁₀	3.34







FIGURE 10: Actual and predicted angle by LSTM model in multi-step-ahead prediction configuration compared with RDPI model. (a) I_1 as input signal. (b) I_2 as input signal. (c) I_3 as input signal. (d) I_4 as input signal. (e) I_5 as input signal.

TABLE 6: Comparison of prediction errors for LSTM and RDPI models across 5 input types.

Input set		I_1	I_2	I_3	I_4	I_5
RMSE (deg)	RDPI model	0.8501	0.9313	0.5734	0.8845	0.8041
	LSTM model	0.1807	0.3509	0.2269	0.2889	0.1518

The model's performance accuracy is lower when making multi-step-ahead predictions compared to one-stepahead predictions, as seen in Figure 9. This outcome is expected because error accumulates over time when the model recursively predicts multiple steps into the future. Table 4 shows the model's RMSE values for multistep predictions across the five input groups I_1 to I_5 .

4. Comparison of the Proposed Model with RDPI Model

The rate-dependent Prandtl-Ishlinskii (RDPI) model is often used to model hysteresis behavior in SMAs and other smart actuators [11, 30]. This model can also incorporate the effect of excitation frequency in its equations. In this study, we utilized the same data that was previously used to train an NN to find appropriate coefficients for the Prandtl-Ishlinskii model as presented in a 2019 paper [11]. The obtained coefficients for the constructed model are shown in Table 5. Furthermore, the results of this RDPI model for the five previously mentioned test data groups are compared to the proposed model in multi-step-ahead prediction configuration in Figure 10, and the results are presented in Table 6.

The results demonstrate that the RDPI model is unable to accurately capture the system's response at peak values, as seen in Figure 10. In contrast, the proposed LSTM model successfully models these points in its multi-step-ahead prediction mode, providing significantly more accurate results. Furthermore, when comparing the errors of the two models in Table 6, it is evident that the LSTM model's errors are approximately 70% lower on average than those of the RDPI model.

5. Conclusion

Modeling the behavior of SMAs is challenging due to their nonlinear dynamics and hysteresis. This research is aimed at creating a model for predicting the pulley rotational angle in an SMA wire-driven rotational actuator. To capture the hysteresis behavior of the SMA, which depends on current and previous inputs to the SMA wire, we employ an LSTM recurrent neural network capable of retaining previous input information.

The LSTM model developed in this research could effectively predict the nonlinear hysteretic behavior of the SMA wire actuator with high accuracy. The model takes as input the SMA wire voltages and pulley angles at the current and two previous time steps and predicts the pulley angle at the next time step. In online configuration where encoder data is available, the LSTM model generates accurate one-stepahead predictions. In offline mode without live encoder data, the LSTM model uses its own predictions as inputs for subsequent time steps. In this configuration, the results of the LSTM model were compared to a rate-dependent Prandtl-Ishlinskii model, highlighting the LSTM's superior accuracy. The success of the LSTM model in accurately capturing the complex hysteretic dynamics of the SMA wire actuator is the key outcome of this research.

Data Availability

The LSTM network prediction results used to support the findings of this study are available from the corresponding author upon request.

Additional Points

Declaration of Generative AI and AI-Assisted Technologies in the Writing Process. During the preparation of this work, the author used Chat GPT-4 in order to improve clarity, grammar, and readability. After using this tool, the author reviewed and edited the content as needed and takes full responsibility for the content of the publication.

Conflicts of Interest

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

References

- G. B. Kauffman and I. Mayo, "The story of nitinol: the serendipitous discovery of the memory metal and its applications," *The Chemical Educator*, vol. 2, no. 2, pp. 1–21, 1997.
- [2] L. G. Machado and M. A. Savi, "Medical applications of shape memory alloys," *Brazilian Journal of Medical and Biological Research*, vol. 36, no. 6, pp. 683–691, 2003.
- [3] G. Song, N. Ma, and H. N. Li, "Applications of shape memory alloys in civil structures," *Engineering Structures*, vol. 28, no. 9, pp. 1266–1274, 2006.
- [4] K. Tanaka, S. Kobayashi, and Y. Sato, "Thermomechanics of transformation pseudoelasticity and shape memory effect in alloys," *International Journal of Plasticity*, vol. 2, no. 1, pp. 59–72, 1986.
- [5] C. Liang and C. A. Rogers, "One-dimensional thermomechanical constitutive relations for shape memory materials," *Journal of Intelligent Material Systems and Structures*, vol. 1, no. 2, pp. 207–234, 1990.
- [6] H. Prahlad and I. Chopra, "Comparative evaluation of shape memory alloy constitutive models with experimental data," *Journal of Intelligent Material Systems and Structures*, vol. 12, no. 6, pp. 383–395, 2001.
- [7] L. C. Brinson, "One-dimensional constitutive behavior of shape memory alloys: thermomechanical derivation with non-constant material functions and redefined martensite internal variable," *Journal of Intelligent Material Systems and Structures*, vol. 4, no. 2, pp. 229–242, 1993.
- [8] J. H. Chung, J. S. Heo, and J. J. Lee, "Implementation strategy for the dual transformation region in the Brinson SMA constitutive model," *Smart Materials and Structures*, vol. 16, no. 1, article N1, 2006.
- [9] K. K. Ahn and N. B. Kha, "Modeling and control of shape memory alloy actuators using Preisach model, genetic algorithm and fuzzy logic," *Mechatronics*, vol. 18, no. 3, pp. 141– 152, 2008.

- [10] M. Zhou, S. He, B. Hu, and Q. Zhang, "Modified KP model for hysteresis of magnetic shape memory alloy actuator," *IETE Technical Review*, vol. 32, no. 1, pp. 29–36, 2015.
- [11] S. Shakiba, A. Yousefi-Koma, M. Jokar, M. R. Zakerzadeh, and H. Basaeri, "Modeling and characterization of the shape memory alloy-based morphing wing behavior using proposed rate-dependent Prandtl-Ishlinskii models," *Proceedings* of the Institution of Mechanical Engineers, Part I: Journal of Systems and Control Engineering, vol. 234, no. 4, pp. 550– 565, 2020.
- [12] M. Al Janaideh, S. Rakheja, and C. Y. Su, "A generalized Prandtl–Ishlinskii model for characterizing the hysteresis and saturation nonlinearities of smart actuators," *Smart Materials and Structures*, vol. 18, no. 4, article 045001, 2009.
- [13] S. M. Dutta and F. H. Ghorbel, "Differential hysteresis modeling of a shape memory alloy wire actuator," *IEEE/ASME Transactions on Mechatronics*, vol. 10, no. 2, pp. 189–197, 2005.
- [14] T. Davenport and R. Kalakota, "The potential for artificial intelligence in healthcare," *Future Healthcare Journal*, vol. 6, no. 2, pp. 94–98, 2019.
- [15] H. Habehh and S. Gohel, "Machine learning in healthcare," *Current Genomics*, vol. 22, no. 4, pp. 291–300, 2021.
- [16] X. Qu, F. Zhao, L. Gao, and Z. Zhang, "The application of machine learning regression algorithms and feature engineering in practical application," in 2022 10th International Conference on Information Systems and Computing Technology (ISCTech), pp. 259–263, Guilin, China, 2022.
- [17] J. Ma, D. Lei, Z. Ren, C. Tan, D. Xia, and H. Guo, "Automated machine learning-based landslide susceptibility mapping for the Three Gorges Reservoir area, China," *Mathematical Geosciences*, vol. 55, pp. 1–36, 2023.
- [18] Z. Liu, J. Ma, D. Xia et al., "Toward the reliable prediction of reservoir landslide displacement using earthworm optimization algorithm-optimized support vector regression (EOA-SVR)," *Natural Hazards*, vol. 119, pp. 1–24, 2023.
- [19] Y. Long, W. Li, R. Huang, Q. Xu, B. Yu, and G. Liu, "A comparative study of supervised classification methods for investigating landslide evolution in the Mianyuan River Basin, China," *Journal of Earth Science*, vol. 34, no. 2, pp. 316–329, 2023.
- [20] Z. Jia, J. Peng, Q. Lu et al., "A comprehensive method for the risk assessment of ground fissures: case study of the eastern Weihe Basin," *Journal of Earth Science*, vol. 34, no. 6, pp. 1892–1907, 2023.
- [21] G. Song, V. Chaudhry, and C. Batur, "A neural network inverse model for a shape memory alloy wire actuator," *Journal of Intelligent Material Systems and Structures*, vol. 14, no. 6, pp. 371–377, 2003.
- [22] E. Asua, J. Feutchwanger, A. García-Arribas, and V. Etxebarria, "Sensorless control of SMA-based actuators using neural networks," *Journal of Intelligent Material Systems and Structures*, vol. 21, no. 18, pp. 1809–1818, 2010.
- [23] M. R. Zakerzadeh, M. Firouzi, H. Sayyaadi, and S. B. Shouraki, "Hysteresis nonlinearity identification using new Preisach model-based artificial neural network approach," *Journal of Applied Mathematics*, vol. 2011, Article ID 458768, 22 pages, 2011.
- [24] H. Wang and G. Song, "Innovative NARX recurrent neural network model for ultra-thin shape memory alloy wire," *Neurocomputing*, vol. 134, pp. 289–295, 2014.

- [25] B. B. Sul, C. S. Subudhi, and K. Dhanalakshmi, "Neural network based displacement modeling of shape memory alloy spring actuator," in 2018 IEEE SENSORS, pp. 1–4, New Delhi, India, 2018.
- [26] A. K. Sheshadri, S. Singh, B. A. Botre et al., "AI models for prediction of displacement and temperature in shape memory alloy (SMA) wire," *AIP Conference Proceedings*, vol. 2335, no. 1, 2021.
- [27] H. N. Bhargaw, S. Singh, B. A. Botre, S. A. Akbar, S. A. R. Hashmi, and P. Sinha, "Deep neural network-based physicsinspired model of self-sensing displacement estimation for antagonistic shape memory alloy actuator," *IEEE Sensors Journal*, vol. 22, no. 4, pp. 3254–3262, 2022.
- [28] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, 1997.
- [29] Y. Bengio, P. Simard, and P. Frasconi, "Learning Long-term dependencies with gradient descent is difficult," *IEEE Transactions on Neural Networks*, vol. 5, no. 2, pp. 157–166, 1994.
- [30] O. Aljanaideh, M. Al Janaideh, S. Rakheja, and C. Y. Su, "Compensation of rate-dependent hysteresis nonlinearities in a magnetostrictive actuator using an inverse Prandtl–Ishlinskii model," *Smart Materials and Structures*, vol. 22, no. 2, article 025027, 2013.