ADVANCED INTELLIGENT SYSTEMS Open Access
www.advintellsyst.com

A Neural Network-Based Self-Sensing Embedded Position Control System for Shape Memory Alloy Wire Actuators

Krunal Koshiya, Gianluca Rizzello, and Paul Motzki*

Shape memory alloy (SMA) wire transducers undergo a phase change in their crystal lattice when heated above a certain transformation temperature, which results in an actuation strain on the order of 4–6% as well as a change in the material electrical properties. The latter opens up the possibility of performing self-sensing, namely estimation of the wire mechanical deformation during actuation based on electrical measurements only. However, accurate SMA self-sensing is challenging because the electromechanical and thermal characteristics are generally hysteretic and depend on various factors like actuation frequency, biasing mechanism, external load, and electrical heating strategy. To practically address this issue, in this work, an artificial-intelligence-based approach is proposed for SMA self-sensing. The proposed neural network consists of a combination of recurrent and simple neuron types to reconstruct the position of an SMA actuator in real-time, using applied electrical power and SMA wire's electrical resistance as inputs. After training and validation, the neural network is implemented on a microcontroller-based embedded system to develop a self-sensing closed-loop position control system. The accuracy of the self-sensing control scheme is validated against a sensor-based control solution, showing root mean squared errors of 0.028 and 0.021 mm respectively over a displacement range of 2 mm.

1. Introduction

Shape memory alloy (SMA) transducers are smart metallic materials that can return to their pre-defined shape when heated by a thermal stimulus or loaded with a mechanical force, a behavior which is commonly referred to as *shape memory effect*.[1] When used as actuators, SMAs are commonly shaped as wires and pretensioned with a mechanical biasing mechanism, e.g., a spring. The application of an electric current to the SMA-wire results in an increase in temperature which, in turn, causes a contraction in length of about 4–6% as a result of a phase transformation occurring in the material crystal lattice. When the current is removed, and the wire returns to lower temperature due to natural convection, the wire restores its original length because of the preload force provided by the biasing mechanism.[1] Thanks to their compact size, noise-free actuation, biocompatibility, high force-to-weight ratio, and vibration damping capability, SMAs are suitable for a wide range of applications, including robotic hand grippers,[2] soft robots,[3] bionic exoskeletons,[4] wearables,[5] adaptive aerodynamics structures,[6] and self-expanding stents.[7]

The electromechanical characteristic curve of SMA wires is in general strongly nonlinear and hysteretic, making their accurate position/force control highly challenging.[8] To effectively implement closed-loop control strategies aimed at compensating such hysteresis, a reliable measurement of the actuator's mechanical output must be available.[9] In conventional closed-loop systems, this feedback signal is measured by an external electromechanical transducer, which unavoidably increases the overall system size, complexity, cost, and encumbrance. A potential alternative offered by SMA technology is to use the transducer itself as a sensor while it is being operated as actuator, a mode referred to as *self-sensing*. Indeed, during phase transformation, the resistivity of the SMA-wire changes depending on its geometry, the material phase composition, and temperature.[10] This feature opens up the possibility of estimating the SMA actuator mechanical output (e.g., displacement, force) based on real-time processing of its electrical signals (e.g., voltage and current) measured during actuation.[11] The self-sensed displacement can then be used as feedback signal in a closed-loop system, leading to a so-called *sensorless* control architecture, which is more integrated, more affordable, and smaller in size compared with its sensorized counterpart.[12]

To effectively achieve self-sensing in SMA actuators, the relationship between the material electrical quantities and the corresponding mechanical output must be properly characterized by means of a model. Accurate modeling of the electro-mechanical characteristics of SMAs for self-sensing applications is in general a hard task, since this curve is generally nonlinear, hysteretic, and non-monotonic and depends on several external factors such

K. Koshiya, G. Rizzello, P. Motzki
The Department of Systems Engineering
Saarland University
66123 Saarbrücken, Germany
E-mail: krunal.koshiya@uni-saarland.de

K. Koshiya, P. Motzki
ZeMA- Center for Mechatronics and Automation Technology
66121 Saarbrücken, Germany

The ORCID identification number(s) for the author(s) of this article can be found under https://doi.org/10.1002/aisy.202501204.

DOI: 10.1002/aisy.202501204

as biasing mechanism, additional external loads, actuation frequency, ambient temperature, and electrical activation strategy. In most research works dealing with SMA self-sensing, the wire electrical resistance is used as main electrical parameter to estimate the strain of the wire, e.g., based on a polynomial fit of the resistance-strain characteristics[13] or via more advanced Kalman filter approaches.[14] Researchers have also emphasized the role of self-sensing in enhancing the performance of control strategies in SMA actuated systems.[12,15] In ref. [16], the authors manipulated the tensile force applied on the wire in order to minimize the hysteresis width of the resistance-strain characteristic to accurately control the position of an SMA actuator via self-sensing. Neural network-based control strategies are presented to address nonlinearity and hysteresis, e.g., a gate recurrent neural network-based model is presented to design a feedforward controller to compensate the nonlinearity and hysteresis that improves the tracking performance of piezoelectric actuators.[17] In ref. [18], Wang et al. highlighted the importance of iterative compensation using neural network-based modeling in adaptive robust control strategies in addressing challenges associated with nonlinearities, hysteresis, and external disturbances.

In the self-sensing characteristics of SMA actuated systems, outer hysteresis loop refers to the hysteretic loop obtained when the SMA wire transforms from full martensite phase to full austenite phase and transforms back from full austenite phase to full martensite phase. Any intermediate transformations forms the minor/inner hysteresis loops as shown in **Figure 1**. These minor hysteresis loops are addressed in ref. [19] using LSTM-based neural networks, in ref. [20] by a PWM-based self-sensing feedback controller with inverse hysteresis compensation, in ref. [14] by an unscented Kalman filter, and in ref. [21] by a fuzzy-PID controller. It is remarked how artificial intelligence provides a potential means to improve the accuracy of self-sensing estimation in SMA,[22] for example, a machine learning-based approach for self-sensing using random forest regression,[23] and an artificial neural network-based self-sensing by employing long-short-term-memory neurons.[19,24] Based on those works, it can be seen that neural networks allow to achieve better accuracy in the presence of hysteresis, overcoming the performance of standard polynomial regressions or other simple methods. However, present studies are limited to describing either outer hysteresis loops only, or minor loops obtained under well-defined and simple

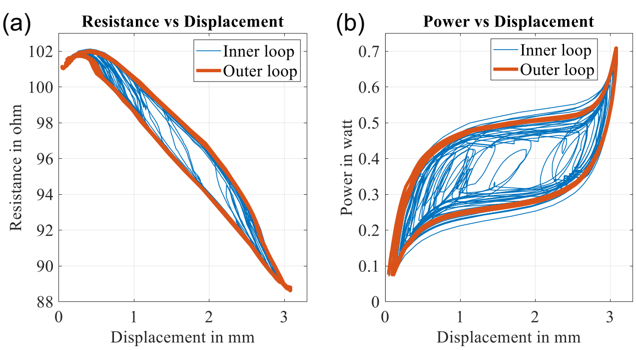

Figure 1. Hysteretic correlations between: a) resistance and displacement and b) power and displacement.

loading trajectories. Nevertheless, achieving accurate self-sensing-based feedback control in SMA actuators subjected to complex input waveform, characterized by random amplitudes and frequencies, still remains an open problem to the best of our knowledge. The complexity of this task stems from the need of accurately modeling the hysteresis non-monotonicity, the internal hysteresis loops, and their dependency on the driving frequency. Even though different types of neural networks are available nowadays, selecting the best model for SMA self-sensing applications requires finding the optimal trade-off between accuracy and real-time computational complexity.

A preliminary investigation on self-sensing-based control of SMA was presented in our previous work,[25] where the feasibility of neural network-based self-sensing control of SMA was demonstrated by tracking the hysteresis outer loop. However, aspects related to the generalizability of the neural network approach to arbitrary inner-loop control trajectories or to different types of SMA actuators, as well as the systematic investigation of the performance-complexity trade-off of the self-sensing method, were not studied in ref. [25]. In this article, we extend the preliminary investigation in ref. [25] by addressing all the abovementioned points, while also making sure that the real-time implementation requirements are met on a microcontroller-based embedded control systems operating at 1 kHz. Given the conflicting objectives of accuracy (favoring complex networks) and real-time feasibility (favoring simpler designs), the investigation focuses on relatively simple type of neurons: feedforward neurons as they require minimum computation power and simple recurrent neurons as they encode temporal dependencies while being relatively simple. It is remarked that more complex recurrent neurons, such as LSTM[19] or GRU,[17] are also commonly used for modeling the temporal dependencies. Since these types of neurons require more computation time due to increased complexity, they are not considered in this work. The input quantities used to reconstruct the SMA displacement via neural networks are the wire electrical resistance and driving input power, both computed in real time via voltage and current measurements. As a further contribution, we prove that using the input power in addition to resistance, which is the only self-sensing signal used in most of the works from the literature,[13,14,16] has the advantage of addressing the non-monotonicity of the resistance-strain curve, leading to better accuracy in the full actuation range. In summary, our work contributes to the state of the art on SMA self-sensing via the following points: 1) We perform systematic parameter studies aimed at finding the number of neurons ensuring the best accuracy compatibly with real-time requirements; 2) we quantitatively assess the self-sensing accuracy improvement when considering both resistance and power as input instead of resistance only; 3) we propose a novel neural network structure for SMA displacement estimation, which differs from other state of the art solution in the combined use of resistance and power data as input as well as the adoption of recurrent neurons specifically in the required hidden layer; 4) we prove generality of the approach by validating it experimentally on three different SMA wire actuators; and 5) we assess the capability of the new neural network in enabling real-time self-sensing-based position control on a microcontroller-based embedded system, by conducting experimental studies to compare the performance of the self-sensing-based control with a sensor-based one.

The remainder of this work is organized as follows. In Section 2, the working principle of the SMA actuated system is described, along with the experimental system under investigation. Section 3 introduces the optimized neural network architecture that is used for SMA self-sensing, and presents its experimental validation. Section 4 systematically compares the sensor-based and self-sensing-based closed-loop-control. Finally, Section 5 provides concluding remarks.

2. SMA-Wire Actuator System

This section begins with a brief overview of the actuation and sensing principles of SMA wires. Subsequently, the experimental devices used to validate the proposed self-sensing method are described.

2.1. SMA Operating Principle

The crystal lattice of SMAs is stable at mainly two solid-state phases, i.e., martensite (stable at lower temperature) and austenite (stable at higher temperature). In martensite phase, each crystal can also have a different orientation, called variant. Based on the orientation of each variant, SMA wires can be in twinned martensite phase (favored in a load-free condition) and detwinned martensite phase (favored when an external load induces mechanical stress).[26] In actuator applications, SMAs are oftentimes shaped as wires and are preloaded with a mechanical biasing mechanism, which stabilizes the material in detwinned martensite phase at room temperature. When an electric current is applied to the wire, the increase in temperature resulting from Joule heating causes the wire to transform from detwinned martensite to austenite. As a result of this phase transformation, the wire length is reduced because the austenite phase has a cubic space-centered crystal structure, making it shorter than the detwinned martensite phase which has instead a unidirectional monoclinic crystal structure.

An experimental setup comprising a spring-loaded SMA actuated system is developed for the practical evaluation of self-sensing-based closed-loop control. When an SMA wire is connected to a linear tensile spring as shown in **Figure 2**a, the wire is under tensile stress against the spring. In this configuration, the spring and the SMA wire exert an equal amount of force on each other at a stable equilibrium point, and the wire is in the detwinned martensite phase at room temperature. If the SMA-wire is heated above its austenite finish temperature, phase transformation from martensite to austenite occurs. As a result, the stiffness of the SMA wire increases, thus altering the stress-strain curve of the wire, as illustrated in Figure 2b.[27] Due to the increased stiffness of the SMA wire, it pulls the spring in the direction of the wire, causing the equilibrium point to shift in the direction of SMA wire. Once the current is no longer applied, the wire cools down to room temperature and reverts to the martensite phase, returning the system to its initial equilibrium point.

2.2. Electromechanical System

An SMA wire (Flexinol) with a 50 μm diameter and 70 °C austenite start temperature, supplied by Dynalloy,[28] is utilized in the experimental setup. The schematic representation in Figure 2a illustrates an SMA-wire actuator configuration in which a linear spring of 0.05 N mm^{-1} stiffness functions as a bias system. As shown in Figure 1a, the SMA wire is formed into a U-shape to enable both wire terminals to be on the fixed end. The length of the SMA wire is calculated as follows.

$$L_{\mathrm{SMA}} = 2l_{\mathrm{SMA}} + \pi r \tag{1}$$

where L_{SMA} is the total wire length, l_{SMA} is the distance between the terminal and the starting point of the curvature at which the wire is bent to form a U-shape, and r is the radius of curvature. The schematic diagram in Figure 2a also demonstrates that increasing the total length between the two fixed endpoints, denoted as D_{Total}, results in an increased force exerted by the spring and SMA wire on each other. This is crucial to establish an initial tensile stress on the SMA wire, ensuring that it is in the detwinned martensite phase when no electrical current is applied.

In the experimental setup illustrated in **Figure 3**, the length of the SMA wire is known when it is in complete austenite phase, and it is equal to 200 mm. An initial tensile stress of 180 MPa is maintained to ensure the complete transformation of the wire to the detwinned martensite phase at room temperature. To set the initial tensile stress of 180 MPa, the initial tensile force is calculated as follows.

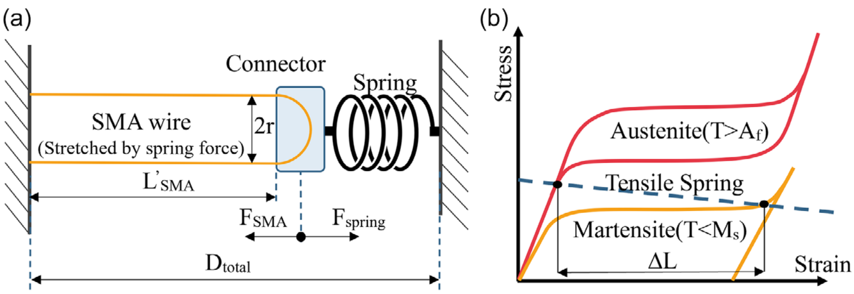

Figure 2. Working principle of a spring-loaded SMA actuated system. a) Schematic diagram of the experimental setup and b) stress-strain characteristics of the SMA-wire during actuation.

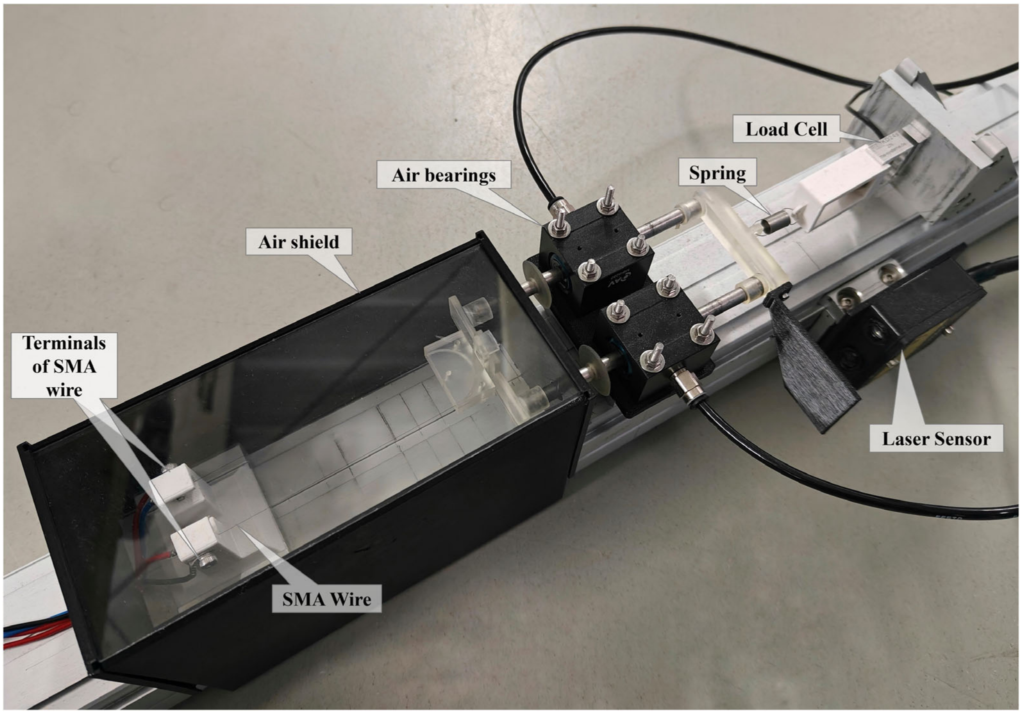

Figure 3. Experimental setup of an SMA-wire actuator configuration with a linear tensile spring as a bias system.

$$F_{\text{initial}} = \sigma_{\text{initial}} * A = \sigma_{\text{initial}} * \pi(r)^2 = (180 * 10^6) * \pi(25 * 10^{-6})^2 = 0.35\text{N} \tag{2}$$

where F_{initial} is the initial tensile force, σ_{initial} is intital tensile stress (180 MPa), A is the cross sectional area of the wire, and r is the radius of the wire (25 μm). The required initial tensile force (0.35 N) is measured using a force sensor KD24s 2 N,[29] which can measure up to 2 N force with 0.1% accuracy. The known austenite length and the fixed starting position are essential to ensure the repeatability of the experimental results. Air bearings are used to provide frictionless movement. The experiments have been performed in a laboratory where ambient temperature is set to 23 °C with allowed deviation of ±1 °C. It is important to maintain the ambient temperature at 23 °C ± 1 °C because the phase transformation of SMA is temperature sensitive, and further change in the ambient temperature might affect its hysteretic characteristics significantly. To minimize the impact of uncontrollable parameters such as air velocity around the wire, an air shield is installed on the setup to isolate the SMA-wire from its surroundings. An optoNCDT 1607–10 laser sensor by micro-epsilon,[30] which has measurement resolution of 3 μ m and measuring range of 10 mm, is used to measure the displacement.

2.3. Embedded System

Along with the electromechanical system, an embedded system is developed to perform various experiments and to collect data on the setup depicted in Figure 3, based on microcontroller STM32H743ZI2. The microcontroller-based embedded system makes the self-sensing SMA actuators easy to adopt in compact SMA actuated systems. The block diagram of the embedded system is presented in **Figure 4**.

It serves multiple purposes, such as gathering characterization data, implementing a trained neural network for self-sensing, and controlling the position of the SMA actuator. Based on the algorithm of specific experimental requirements, an analog voltage signal is produced by the microcontroller STM32H743ZI2 using DAC and is passed to a current source. The current source produces an electrical current based on the input voltage. The regulated electrical current is applied to the SMA wire for heating the wire. As the wire is heated above its austenite finish temperature, it stretches the spring, and the displacement is measured using a laser sensor. The voltage across the SMA wire as well as the current provided by the current source are measured. The outputs of the laser sensor, measured voltage, and measured current signals are converted into a 0–3 V range, which can be read via ADC using the microcontroller. The ADC readings are stored in a circular buffer for use in the algorithm and for transmission to a personal computer via USART for offline analysis.

For collecting the characterization data, an experiment is performed in which an electrical input current of random amplitude between 0 and 90 mA and random frequency between 20 and 200 mHz is applied to the SMA wire. A sampling rate of 1 kHz is chosen for signal output generation and measurements. The displacement D achieved by the actuator is measured using a laser sensor, while applied electrical power P and electrical resistance R are calculated from measured current i and voltage v using

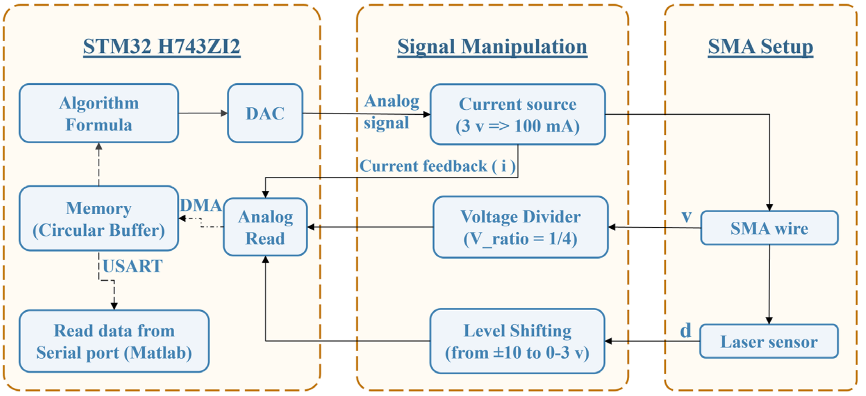

Figure 4. Block diagram of the microcontroller-based embedded system.

$$P = vi, \quad R = \frac{v}{i} \tag{3}$$

After collecting the characterization data, a half-Gaussian filter is implemented to reduce the measurement noise. The half-Gaussian filter is implemented as follows

$$y_t = \frac{\sum_{i=n}^{0} w_i u_{t+i}}{\sum_{i=n}^{0} w_i} \tag{4}$$

$$w_i = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{S_i^2}{2\sigma^2}} \tag{5}$$

where $n \leq 0$ is the filter window length, σ represents the standard deviation of the Gaussian distribution, y_t is the filtered value at time t, u_{t+i} is the i^{th} input value from the filter window at time t, w_i is the weight for i^{th} input value, and S_i is the sample number in the filter window. Negative sample numbers indicate the previous values that are used for filtering the measured values. Positive sample numbers are not used in the half-Gaussian filter during the real-time implementation of the algorithms, as this would violate causality of implementation. The characterization data of the SMA actuated systems, consisting of applied electrical power and resistance, as well as resulting displacement are filtered using the half-Gaussian filter as presented in **Figure 5**.

It is practically observed that the presented SMA-wire-spring actuator exhibits a hysteretic correlation between resistance and displacement as well as power and displacement, as shown in Figure 1. It is critical to note that slower actuation frequencies lead to a reduction in the width of the hysteresis loop in the power-displacement curve. Additionally, the intermediate transformation of the SMA-wire to the austenite phase results in the formation of inner hysteresis loops, as shown in Figure 1, which complicate the overall modeling. As illustrated in Figure 1b, several blue curves appear outside the outer loop indicated by the red curves in power-displacement plot. These blue curves correspond to inner hysteresis loops formed during intermediate phase transformations, similar to the other inner loops. However, their hysteresis width is larger due to the slower actuation frequencies. For the training, validation and testing of neural network architectures, the electrical current signal of random amplitude and frequency is applied to the SMA wire to ensure the model's robustness and to avoid the risk of overfitting.

3. Artificial Neural Networks

After describing the neural network analysis, this section presents experimental validation based on the data collected on the experimental setup.

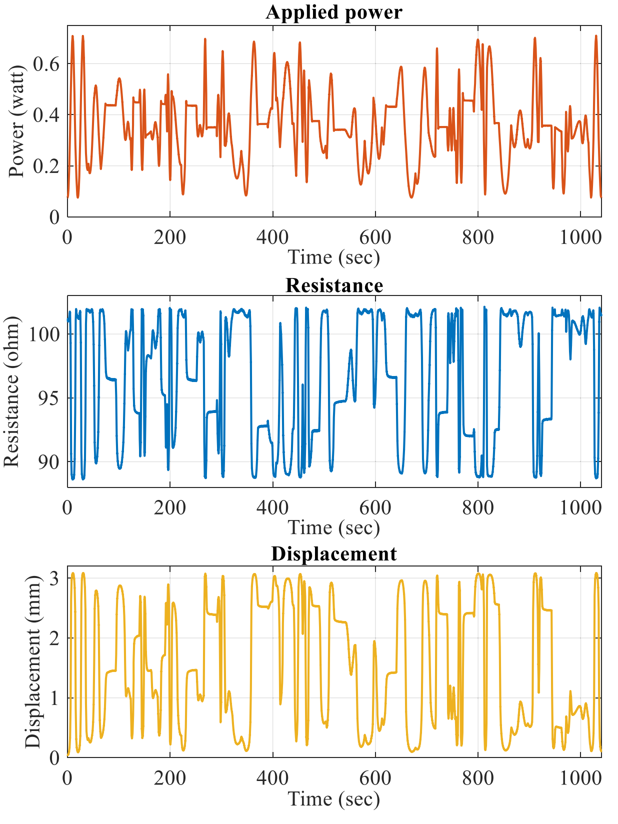

Figure 5. Characterization data of a spring-loaded SMA-wire actuator.

Table 1. Analysis of different neural network architectures for self-sensing in SMA actuators.

NN	Input	H1 Neurons [Sigmoid]	H2 Neurons [ReLU]	Training RMSE [mm]	Validation RMSE [mm]	Execution time [mSec]
1	R	8 (Simple)	–	0.167	0.169	0.14
2	R	16 (Simple)	–	0.168	0.169	0.16
3	R	16 (Simple)	8 (Simple)	0.166	0.168	0.19
4	R	32 (Simple)	16 (Simple)	0.163	0.175	0.29
5	R	16 (Recurrent)	8 (Simple)	0.149	0.153	0.86
6	R, P	8 (Simple)	–	0.065	0.070	0.17
7	R, P	16 (Simple)	–	0.063	0.068	0.20
8	R, P	16 (Simple)	8 (Simple)	0.051	0.058	0.23
9	R, P	32 (Simple)	16(Simple)	0.047	0.051	0.31
10	R, P	16 (Recurrent)	8 (Simple)	0.020	0.019	0.94
11	R, P	16 (Recurrent)	8 (Recurrent)	0.023	0.024	1.28

R = electrical resistance, P = electrical power.

3.1. Neural Networks Analysis

An artificial neural network is composed by interconnected nodes, known as neurons, which are organized in various layers connected in series.[31] Each neuron processes input data and passes the result to subsequent layers to produce an output. As reported in **Table 1**, we evaluate various input combinations and neuron counts for each network type, to determine the optimal configuration that achieves maximum accuracy without compromising real-time feasibility. The “Input” column indicates the neural network inputs that are used to estimate the displacement. A single neuron in the output layer gives estimated displacement as an output with ReLU activation function. The “H1 Neurons” column refers to the number and type of neurons used in the first hidden layer with a Sigmoid activation function, while the “H2 Neurons” column refers to those used in the second hidden layer with a ReLU activation function. The “Execution time” column displays the time required to predict displacement using the trained network when implemented on a microcontroller STM32H743ZI2, operating at a clock frequency of 200 MHz. “Training RMSE” and “Validation RMSE” quantifies the performance of each NNs on training and validation data, respectively, by the root mean squared error (RMSE).

Comparing NNs 1 and 2, 3 and 4, 6 and 7, and 8 and 9, it can be observed that an increase in the number of neurons leads to better accuracy at the expense of greater computational power. Similarly, comparing NNs 2 and 3 and 7 and 8 shows that adding more layers also leads to improved accuracy but results in increased computation time. For self-sensing in SMA actuators, the conventional method is based on using the electrical resistance as an input for estimating displacement. However, comparisons between NN 1 and 6, 2 and 7, 3 and 8, 4 and 9, and 5 and 10 indicate that incorporating electrical power, in addition to electrical resistance, significantly improves prediction accuracy. It is known that the resistance of SMAs depends on the material strain as well as temperature and internal state.[10] Prior studies show that electrical power correlates with temperature,[32] and phase composition relates to the mechanical state in SMA actuators,[10] emphasizing the dynamic nature of these relationships. By combining electrical resistance and electrical power readings as inputs, the neural network improves the accuracy of the self-sensing model by learning the correlation between electrical resistance, phase composition, and SMA wire geometry, as well as the correlation between electrical power and temperature. The use of recurrent neurons further allows learning these characteristics in a dynamic fashion. This result is achieved with a data-driven approach, to address the lack of direct internal state measurements (e.g., phase, temperature) affecting hysteresis as well as the strong material nonlinearity. In the performance comparison between NN 5 and NN 10 as depicted in **Figure 6**, it is evident that resistance- and power-based neural network models, like NN 10, have the capability to account for the nonmonotonicity of hysteresis, unlike the sole resistance-based model, such as NN 5.

As the hysteresis inherently involves memory of past states, modeling the hysteresis with a neural network involves capturing the dependency on the history of past inputs. The recurrent neurons are designed to handle sequential data because they maintain a form of memory, which allows them to capture dependencies and relationships over time. This feature makes

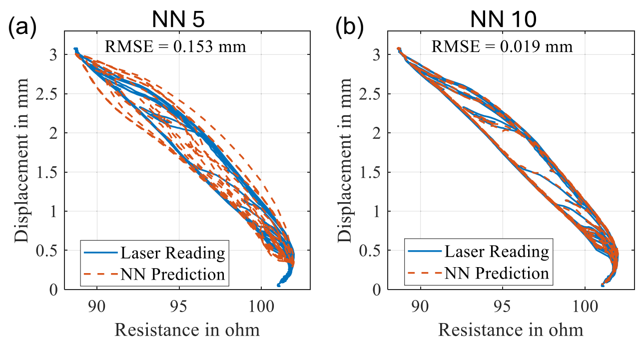

Figure 6. Neural network performance on modeling resistance-displacement characteristic when a) only resistance is used as input (NN5) and b) resistance and power are used as inputs (NN10).

them particularly well-suited for modeling hysteresis. In standard recurrent neurons, the output of a neuron is fed back and used as neuron input for the subsequent time step. This mechanism enables the neuron to retain information about past inputs and incorporate the influence of prior states into the current computations, practically implementing a memory. Recurrent neurons, however, require more computation power than conventional ones. In the proposed NN model, we have used standard recurrent neurons in the first hidden layer to capture the temporal dependencies that are needed for modeling the hysteresis memory effect. The input of a recurrent neuron must be provided in the form of a time series sequence. The input sequence is created by considering the 10 previous values of the inputs with a 100 millisecond delay between consecutive inputs. Therefore, the sequence window covers the previous 1 s of temporal data that is used by recurrent neurons to model the temporal dependencies. From the comparison between NNs 8 and 10, it is evident that replacing simple neurons with recurrent neurons in the first hidden layer reduces RMSE by 60.8% and 67.24% in training and validation, respectively, which proves that recurrent neurons outperform basic neurons in predicting the hysteretic material response. Further comparison between NNs 10 and 11 reveals that as the network's complexity increases, it starts learning the noise patterns as well, instead of focusing on the main patterns, which causes a decrease in prediction accuracy. The required computation time also increases from 0.94 millisecond to 1.28 millisecond. For implementation of an algorithm at 1 KHz sampling frequency, it is necessary that the required computation time for the algorithm implementation is less than 1 millisecond. Based on the above discussion, we select NN 10 to be implemented on a microcontroller running at a sampling frequency of 1 KHz. The proposed neural network structure, illustrated in **Figure 7**, is motivated by the need to find a trade-off between required computation power and accuracy. The structure of our neural network is based on a physical argument regarding the SMA electrical resistance model,[10] even though it is here incorporated in a data-based fashion. In particular, since data on the material's internal state, which directly affects hysteresis (phase, temperature) is generally not available, our neural network attempts to extract the additional information on the material internal state from the combined resistance and electrical power readings, and use them to correct the displacement-resistance model accordingly. Indeed, from earlier modeling works on SMA actuators, it is known that the power is strictly related to the material temperature,[32] while there exists a well-defined relationship between material phase composition and mechanical state for SMA-spring actuators.[33] As those relationships are generally based on a dynamic model, a neural network embedding memory (e.g., recurrent) is expected to perform better than a memoryless one (e.g., feedforward).

A combination of backpropagation algorithm[34] and backpropagation through time algorithm[35] is used for training of the specified neural network, NN 10, in which the recurrent layer is followed by simple layers. The mean squared error using predicted output $\hat{y}$ and actual output y is calculated by

$$L = \frac{1}{2}(\hat{y} - y)^2 \tag{6}$$

for loss computation. All the trainable parameters are updated using ADAM optimizer.[36]

3.2. Performance Evaluation of Neural Network

An input–output dataset is prepared using electrical power and resistance as inputs, and measured displacement as an output, for training the neural network. The characterization data shown in Figure 5 is used for the training of the neural network with ADAM as an optimizer with hyperparameters, $\alpha = 0.001$, $\beta_1 = 0.9$, $\beta_2 = 0.999$, and mean squared error calculated by (6) as a loss function. Apart from the training data, additional measurements are collected for validation purpose. The results for both training and validation data are compared with the laser sensor readings, as shown in **Figure 8**a. To quantify the performance of NN10, the root mean squared error (RMSE) is calculated. The network can estimate the displacement with a RMSE of 0.0196 mm for the training data and 0.0194 mm for the

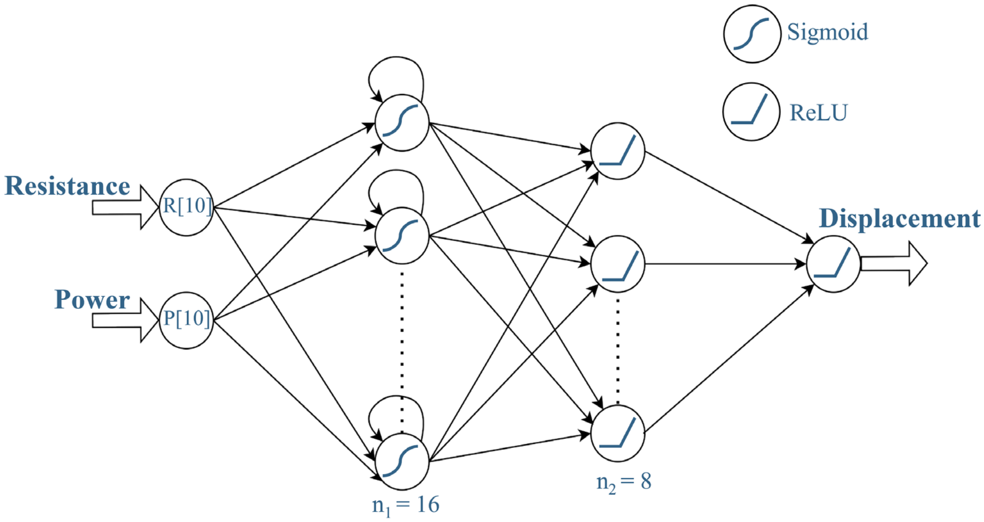

Figure 7. Schematic diagram of the optimized neural network structure for self-sensing in an SMA actuated-system.

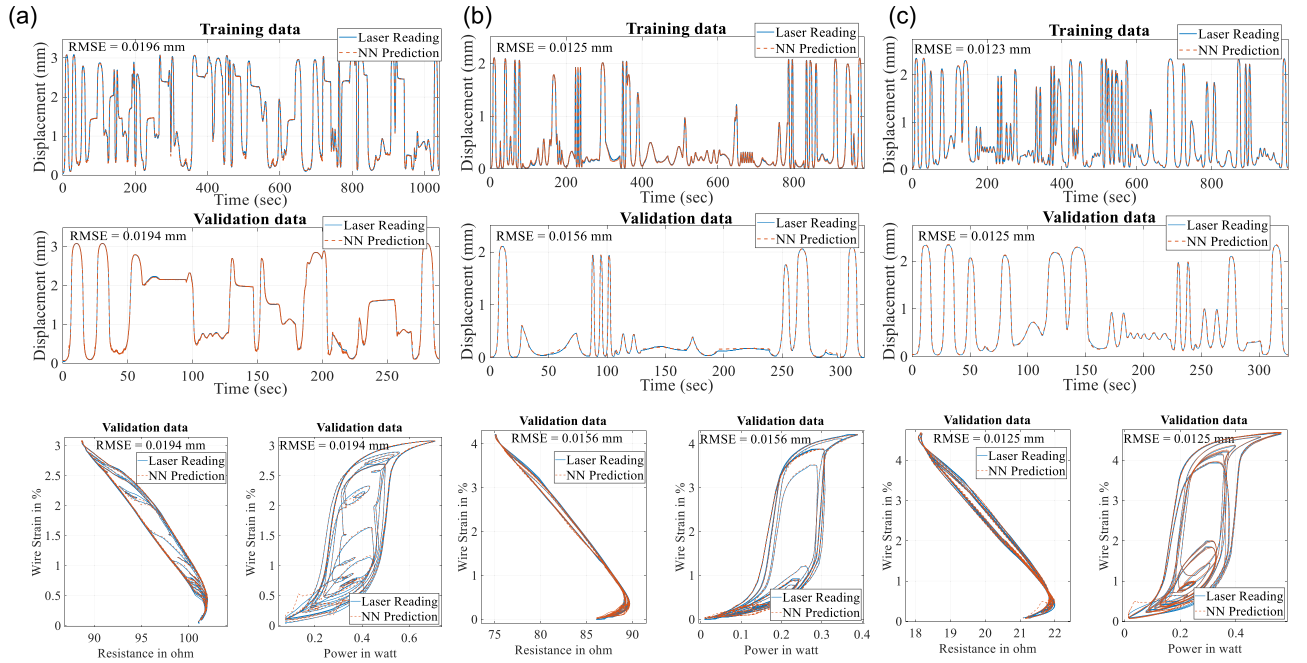

Figure 8. Performance evaluation of the neural network for self-sensing in different SMA wires on training and validation data in comparison to the laser sensor reading. a) Type 1 SMA wire (wire diameter = 50 micrometers, wire length = 200 mm, austenite start temperature = 70 °C, provider = Dynalloy (Flexinol)), b) Type 2 SMA wire (wire diameter = 38 micrometers, wire length = 100 mm, austenite start temperature = 70 °C, provider = Dynalloy (Flexinol)), and c) Type 3 SMA wire (wire diameter = 76 micrometers, wire length = 100 mm, austenite start temperature = 90 °C, provider = SAES getters).

validation data. The experimental results demonstrate that the proposed method achieves self-sensing for arbitrary trajectories exploring SMA inner loops with a precision of 0.0194 mm, corresponding to 0.63% of the maximum displacement. This performance significantly surpasses the best-reported accuracy of 2.5% of the maximum displacement achieved in similar experiments, to the best of our knowledge, using state-of-the-art neural network-based self-sensing methods for SMA actuators by employing LSTM neurons.[19] It is important to note here that the proposed neural network architecture is optimized for real-time implementation at a 1 kHz frequency on a microcontroller-based embedded system, which is challenging to achieve with the state-of-the-art LSTM-based architectures because of their higher computational demands.

Apart from the training and validation on the SMA wire of 50 micrometer diameter (type 1), which is used for all subsequent experiments presented in this paper, we investigated generalization capability of the our self-sensing approach by testing the proposed neural network structure for self-sensing evaluation on multiple SMA wire actuators. For this purpose, two additional experiments are performed using different SMA wire actuators (type 2 and type 3). Characteristics of all wires are reported in the following: 1) Type 1 (wire diameter = 50 micrometer, wire length = 200 mm, austenite start temperature = 70 °C, provider = Dynalloy (Flexinol)), 2) Type 2 (wire diameter = 38 micrometer, wire length = 100 mm, austenite start temperature = 70 °C, provider = Dynalloy (Flexinol)), 3) Type 3 (wire diameter = 76 micrometer, wire length = 100 mm, austenite start temperature = 90 °C, provider = SAES getters).

Characterization data of random amplitude and frequency is collected for each type of the SMA wire. The proposed neural network structure is then trained for each case, using the respective training data with ADAM as an optimizer with hyperparameters, $\alpha = 0.001$, β_1=0.9, β_2=0.999, and mean squared error as loss function. Apart from the training data, additional measurements are collected for validation purpose in each case. As shown in 8, the proposed neural network structure is capable to estimate the displacement for the training and validation data with 0.020 and 0.019 mm, respectively, on type 1; 0.013 and 0.016 mm, respectively, on type 2; and 0.012 and 0.013 mm, respectively, on type 3 SMA actuator. All results are considered satisfactory, and confirm the accuracy and generalizability of the proposed neural network.

Once the network is trained, it can estimate the displacement using applied electrical power and resistance in real-time at 1 kHz frequency using a microcontroller STM32H743ZI2. As shown in **Figure 9**, the trained network can estimate the displacement with 0.0247 mm RMSE when implemented in real-time using the microcontroller-based embedded system. The developed neural network is therefore suitable to generate a feedback signal in a closed-loop control strategy.

4. Sensorless Control Validation

This section validates the use of the developed neural network in a sensorless position control application of an SMA actuator. The goal is to compare a closed-loop position control architecture,

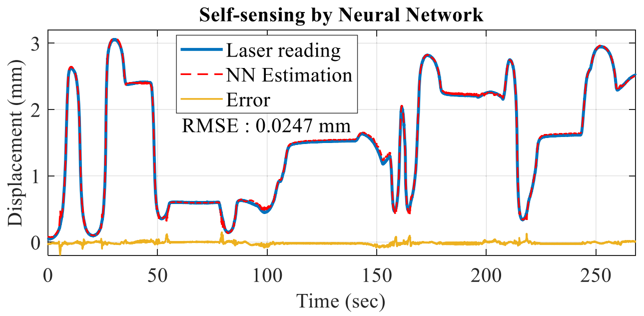

Figure 9. Performance of the trained neural network for self-sensing in a real-time system.

which relies on an accurate laser sensor to acquire the position feedback with a sensorless architecture, which relies on the self-sensing neural network to estimate the position. To perform the experimental comparison among the two closed-loop architectures, the same experimental setup presented in Section 2 is used.

For the SMA-actuated system considered in this work, the relationship between applied current and resulting displacement is highly hysteretic. This strong nonlinearity significantly complicates the design of closed-loop control systems, making it necessary to resort to nonlinear control and/or hysteresis compensation[21,37] whenever high tracking accuracy and/or speed are needed. Nevertheless, since in this work the primary focus is on an accurate self-sensing (rather than control) strategy, we decided to control the SMA actuator position via a simple PI + feedforward law. To this end, the input–output nonlinearity is approximated with a third order polynomial equation passing through the average points of a current-displacement hysteresis curve.[38] Despite their simplicity, PI + feedforward laws represent a popular control architecture in many motion control systems.[39,40]

While the feedforward term performs an approximated inversion of the system dynamics based on a nominal model, the PI attempts at compensating the residual errors, disturbances, and unmodeled dynamics via a feedback action. A block diagram representation of the adopted PI + feedforward law is shown in **Figure 10**. The control input, i.e., electrical current i, is given by

$$i = sat(i_{\mathrm{PI}} + i_{ff}) \tag{7}$$

where i_{PI} and i_{ff} are the PI and feedforward contributions, respectively, while $sat(\cdot)$ denotes a saturation function, which constrains the SMA current within safety bounds at each time, i.e., $i \in [0,90]$ mA. To simplify the implementation and enhance the impact of feedback control, in this work, we let most of the hysteresis compensation being carried out by the PI, while the feedforward only provides a nominal correction term. To this end, the input–output nonlinearity is approximated with a third order polynomial equation passing through the average points of a current-displacement hysteresis curve. In this way, the overall trend of the hysteretic nonlinearity can be compensated without the need to implement a complex hysteresis model. The obtained polynomial fit is then used to implement the following feedforward action, defined as

$$i_{ff} = 7.35D_t^3 - 36.13D_t^2 + 59.41D_t + 29.31 \tag{8}$$

where D_t is the target displacement function. The PI term i_{PI} is implemented as a standard error-driven proportional-integral law with anti-windup, to cope with control input saturation, defined in the z-domain as follows (note that the forward Euler discretization is used).

$$i_{\mathrm{PI}}(z) = \quad k_p e(z) + \tfrac{T_s}{z-1} k_i e(z) - \tfrac{T_s}{z-1} k_a (i_{\mathrm{PI}}(z) + i_{ff}(z) - i(z)) \tag{9}$$

where e is the position tracking error, i is defined as in (8), T_s is the sampling time, while k_p, k_i and and k_a are proportional, integral, and anti-windup gain, respectively.

To control the displacement of the SMA actuator based on a feedback signal from a laser sensor, the overall control law is implemented using a microcontroller, according to the architecture in Figure 10a. In the sensor-based control, the feedback signal is measured by a laser displacement sensor optoNCDT 1607 by micro-epsilon. The PI gains are tuned based on Ziegler-Nichols tuning method and further refined by hand-tuning, resulting in $k_{\mathrm{p}} = 27$, $k_i = 108$, and $k_{\mathrm{a}} = 10.8$. To implement self-sensing-based control, the feedback signal is provided by the neural network trained as in Section 3.2 based on measured

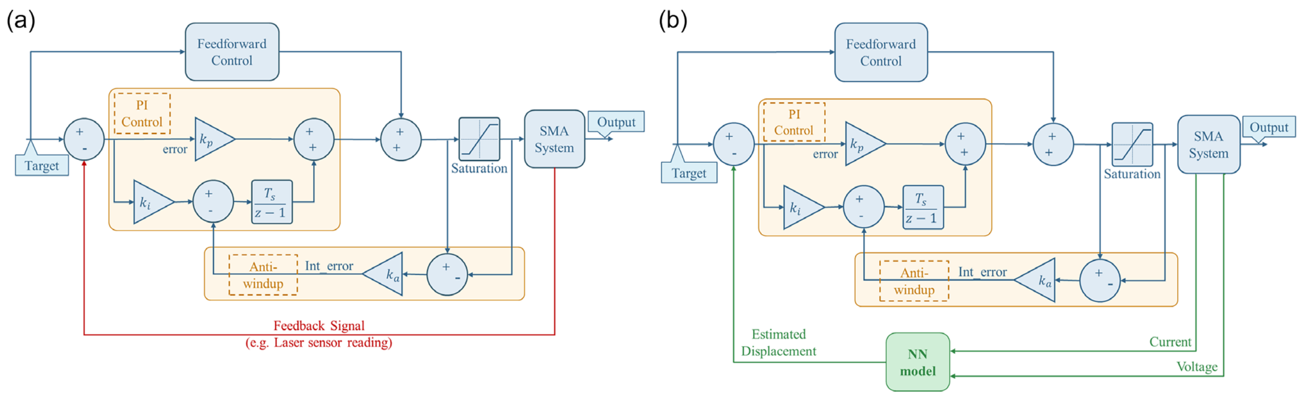

Figure 10. Block diagram of combined control strategies: a) laser sensor-based control and b) self-sensing-based control.

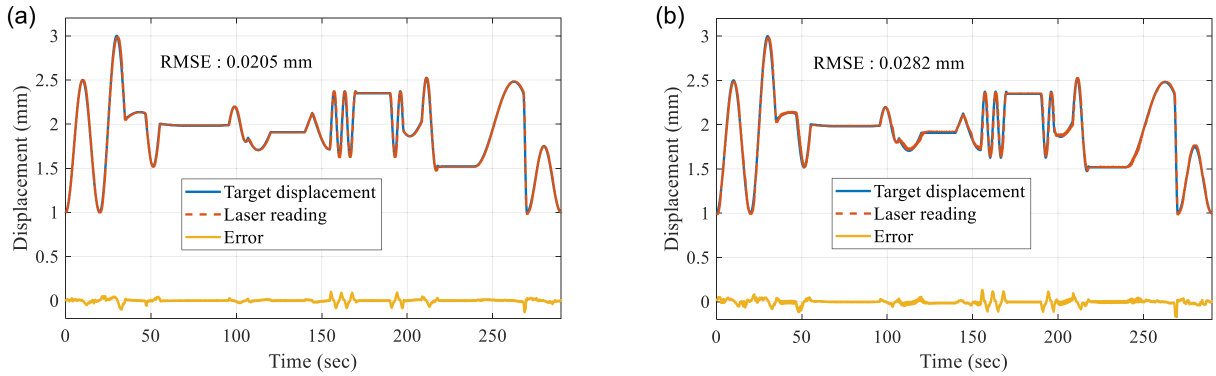

Figure 11. Performance evaluation of a combined control strategy: a) laser sensor-based control (RMSE = 0.0205 mm) and b) self-sensing-based control (RMSE = 0.0282 mm).

voltage and current signals, as shown in Figure 10b. In the implementation of the self-sensing-based control strategy, the output of the laser sensor is replaced by the output of the neural network, and the values of feedback controller gains k_p, k_i and and k_a are kept same as the laser sensor-based control for ease of comparison.

The performance of the laser sensor-based combined position control for the spring-loaded SMA actuated system is presented in **Figure 11**a. The target displacement trajectory is planned as a composition of sinusoidal signals with random amplitude and frequency. This type of signal allows to investigate the performance of the self-sensing-based control strategies in the presence of rate-dependent hysteretic dynamics as well as inner hysteresis loops depending upon intermediate amplitudes. Additionally, the target displacement is also kept constant at non-uniform intervals to examine steady-state errors. The resulting RMSE of the laser sensor-based combined control is 0.0205 mm for target displacement of random amplitude and frequency. The performance of the self-sensing-based closed-loop control system is instead presented in Figure 11b. The resulting RMSE value achieved in this case is 0.0282 mm. Note that in this case, the RMSE is calculated by error based on the measurement of the laser sensor, while the controller operates on the estimated error that is calculated by subtracting estimated displacement from the target displacement. In spite of the fact that the feedback signal in self-sensing-based closed-loop control is solely dependent on a neural network, its accuracy is very close to the one of the laser sensor-based closed-loop control, as illustrated by Figure 11. In **Table 2**, the evaluation of the laser sensor-based control and the self-sensing-based control is reported with the quantitative description of the tracking errors by root mean squared error (RMSE), mean absolute error (MAE), and 95% peak error (PE(95%)). The 95% peak error refers to the maximum deviation of the system's response from the setpoint during the transient phase, excluding the largest 5% of outliers. PE(95%) is commonly used to provide a robust measure of peak error by reducing the influence of extreme or rare spikes in the error. The resulting RMSE, MAE, and PE(95%) achieved in laser sensor-based control strategy is 0.0205 mm, 0.0108 mm, and 0.0497 mm, respectively, and in self-sensing-based control strategy is 0.0282 mm, 0.0177 mm, and 0.0685 mm, respectively. These results effectively validates the proposed neural network-based self-sensing approach.

Table 2. Evaluation of laser sensor-based control and self-sensing-based control results by quantitative description of tracking errors.

Control strategy	RMSE	MAE	PE(95%)
Laser sensor-based	0.0205 mm	0.0108 mm	0.0497 mm
Self-sensing-based	0.0282 mm	0.0177 mm	0.0685 mm

RMSE = root mean squared error, MAE = mean absolute error, PE(95%) = 95% peak error.

5. Conclusion

In this article, we presented a self-sensing embedded position control system for a shape memory alloy actuated system using an artificial neural network. The selection of the neural network is motivated by our best physical understanding of the physical mechanism behind SMA resistance, which is known to depend not only on the material strain but also on its internal state.[10] By supplementing resistance data with electrical power and intrinsic memory, the proposed neural network has all required information and the appropriate internal structure to accurately learn the physical process underlying self-sensing. The proof-of-concept is provided by implementing the self-sensing closed-loop position control architecture on the spring-loaded SMA actuated system using a microcontroller-based embedded system, which performs up to the standard of laser sensor-based closed-loop control systems. By the comparison of root mean squared error in position control based on feedback signal from laser sensor (RMSE of 0.0205 mm) and from self-sensing (RMSE of 0.0282 mm), it is proven that neural network-based self-sensing in SMA actuators can potentially replace the conventionally used external electro-mechanical transducers in a closed-loop control systems.

The presented results motivate the use of a self-sensing closed-loop control systems in development of size-, weight-, and cost-efficient SMA actuated systems. The subsequent steps of this work explicitly address the dependence of ambient temperature on SMA self-sensing. In future work, self-sensing will be used for the accurate position control of complex SMA actuated systems such as continuum robots.

Supporting Information

Supporting Information is available from the authors upon requests.

Acknowledgements

The authors would like to acknowledge financial support of the state of Saarland, Germany for funding this research work within iSMAT project through ERDF - European Regional Development Fund (EFRE).

Open Access funding enabled and organized by Projekt DEAL.

Conflict of Interest

The authors declare no conflict of interest.

Data Availability Statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Keywords

embedded control system, neural network, self-sensing, shape memory alloy actuator

Received: September 27, 2025
Revised: November 23, 2025
Published online: January 30, 2026

[1] M. Cho, S. Kim, Structural Morphing using Two-Way Shape Memory Effect of SMA, *Int. J. Solids Struct.* **2005**, *42*, 1759.
[2] S. M. Jo, H.-S. Yoon, Energy-Efficient Tristable Soft Gripper using Shape Memory Alloy Wires for Gripping Convex and Concave Objects, *J. Mech. Robot.* **2024**, *16*, 021013.
[3] Y. Goergen, G. Rizzello, P. Motzki, Systematic Methodology for an Optimized Design of Shape Memory Alloy-Driven Continuum Robots, *Adv. Eng. Mater.* **2024**, *26*, 2301502.
[4] Q. Xie, Q. Meng, W. Yu, R. Xu, Z. Wu, X. Wang, H. Yu, Design of a Soft Bionic Elbow Exoskeleton based on Shape Memory Alloy Spring Actuators, *Mech. Sci.* **2023**, *14*, 159.
[5] S.-M. Lee, J. Park, A Soft Wearable Exoglove for Rehabilitation Assistance: A Novel Application of Knitted Shape-Memory Alloy as a Flexible Actuator, *Fashion Textiles* **2024**, *11*, 14.
[6] P. Turabimana, J. W. Sohn, S. Choi, Design and Control of a Shape Memory Alloy-based Idle Air Control Actuator for a Mid-Size Passenger Vehicle Application, *Appl. Sci.* **2024**, *14*, 4784.
[7] R.-J. van Geuns, C. Tamburino, J. Fajadet, M. Vrolix, B. Witzenbichler, E. Eeckhout, C. Spaulding, K. Reczuch, A. L. Manna, R. Spaargaren, H. M. García-García, E. Regar, D. Capodanno, G. V. Langenhove, S. Verheye, Self-Expanding versus Balloon-Expandable Stents in Acute Myocardial Infarction: Results from the APPOSITION II Study: Self-Expanding Stents in St-Segment Elevation Myocardial Infarction, *JACC: Cardiovasc. Interv.* **2012**, *5*, 1209.
[8] Y. Xu, L. Guo, D. Pan, H. Zhou, J. Duan, Hysteresis Operator-based Adaptive Kalman Feedforward Control for the Scanning Motion of Hybrid Reluctance Actuators, *ISA Trans.* **2024**, *155*, 463.
[9] T. A. U. Roshan, B. A. D. J. C. K. Basnayake, Y. W. R. Amarasinghe, D. Wijethunge, N. D. Nanayakkara, in *2018 Moratuwa Engineering Research Conf. (MERCon)*, IEEE, Moratuwa, Sri Lanka **2018**, pp. 460–464, https://doi.org/10.1109/MERCon.2018.8421929.
[10] S. J. Furst, S. Seelecke, Modeling and Experimental Characterization of the Stress, Strain, and Resistance of Shape Memory Alloy Actuator Wires with Controlled Power Input, *J. Intell. Mater. Syst. Struct.* **2012**, *23*, 1233.
[11] T. M. Ganapathy, D. Kaliaperumal, M. Banu Sundareswari, in *2023 16th Inter. Conf. on Sensing Technology (ICST)*, IEEE, Hyderabad, India **2023**, pp. 1–5, https://doi.org/10.1109/ICST59744.2023.10460822.
[12] H. Ju, B. Cha, D. Rus, J. Lee, Closed-Loop Soft Robot Control Frameworks with Coordinated Policies based on Reinforcement Learning and Proprioceptive Self-Sensing, *Adv. Funct. Mater.* **2023**, *33*, 2304642.
[13] C.-C. Lan, C.-H. Fan, An Accurate Self-Sensing Method for the Control of Shape Memory Alloy Actuated Flexures, *Sens. Actuat. A: Phys.* **2010**, *163*, 323.
[14] H. Gurung, A. Banerjee, Self-Sensing Shape Memory Alloy Wire Actuator based on Unscented Kalman Filter, *Sens. Actuat. A: Phys.* **2016**, *251*, 258.
[15] D. J. S. Ruth, J.-W. Sohn, K. Dhanalakshmi, S.-B. Choi, Control Aspects of Shape Memory Alloys in Robotics Applications: A Review Over the Last Decade, *Sensors* **2022**, *22*, 4860.
[16] S.-H. Lee, S.-W. Kim, Improved Position Control of Shape Memory Alloy Actuator using the Self-Sensing Model, *Sens. Actuators A: Phys.* **2019**, *297*, 111529.
[17] B. Hou, Ze Wang, C. Hu, Yu Zhu, Gate Recurrent unit Neural Network based High-Precision Feedforward Control for Piezoelectric Nanopositioning Stage, *Eng. Appl. Artif. Intell.* **2024**, *133*, 108126.
[18] Ze Wang, C. Hu, Yu Zhu, S. He, K. Yang, M. Zhang, Neural Network Learning Adaptive Robust Control of an Industrial Linear Motor-Driven Stage with Disturbance Rejection Ability, *IEEE Trans. Ind. Informat* **2017**, *13*, 2172.
[19] S. Mohan, A. Banerjee, LSTM-based Self-Sensing Application of Shape Memory Alloy Wire Actuators under Practical Loading Conditions, *Shape. Mem. Superelasticity* **2024**, *10*, 407.
[20] S.-H. Liu, T.-S. Huang, J.-Y. Yen, Tracking Control of Shape-Memory-Alloy Actuators based on Self-Sensing Feedback and Inverse Hysteresis Compensation, *Sensors* **2009**, *10*, 112.
[21] F. Bai, X. Zhang, D. Xu, Accurate Position Tracking Control of SMAAs based on Low-Complexity Self-Sensing Model and Compound Control Strategy, *IEEE Sens. J.* **2022**, *23*, 2280.
[22] R. Hmede, F. Chapelle, Y. Lapusta, Review of Neural Network Modeling of Shape Memory Alloys, *Sensors* **2022**, *22*, 5610.
[23] K. Scott, M. R. E. U. Shougat, E. Perkins, Robust Self-Sensing Shape Memory Alloy Actuator using a Machine Learning Approach, *Sens. Actuators, A: Phys* **2023**, *354*, 114255, ISSN 0924-4247.
[24] Y.-C. Pei, B.-H. Wang, J.-T. Wu, C. Wang, J.-H. Guan, H. Lu, A Machine Learning Empowered Shape Memory Alloy Gripper with Displacement-Force-Stiffness Self-Sensing, *IEEE Trans. Ind. Electron.* **2022**, *70*, 10385.
[25] K. J. Koshiya, G. Rizzello, P. Motzki, Neural Network-Based Sensorless Control of a Shape Memory Alloy Actuator, ASME 2024 Conf. on Smart Materials, Adaptive Structures and Intelligent Systems, Smart Materials, Adaptive Structures And Intelligent Systems, American Society of Mechanical Engineers, Atlanta, Georgia, USA **2024**, p. V001T03A001.
[26] D. C. Lagoudas, Shape Memory Alloys, *Sci. Bus. Media, LLC* **2008**, *1*, 5.
[27] P. Motzki, Efficient SMA Actuation—Design and Control Concepts, in *Proc. of The 1st Inter. Electronic Conf. on Actuator Technology: Materials, Devices and Applications*, MDPI, Basel, Switzerland **2020**, Vol. *64*, p. 20.

[28] Technical Characteristics of Flexinol Actuator Wires. Dynalloy Inc, **2025**, URL https://dynalloy.com/wp-content/uploads/2025/03/TCF1140.pdf (accessed: November 2025).
[29] *Force Sensor KD24s 2N*. ME systeme, **2025**, URL https://www.me-systeme.de/en/pdf/1719/KD24s_2N (accessed: November 2025).
[30] Non-Contact Laser-Optical Displacement Measurement - Compact PSD sensor optoNCDT 1607-10. Micro-Epsilon, **2025**, URL https://www.micro-epsilon.pl/download/Datasheet_-_optoNCDT_1607_-_en.pdf (accessed: November 2025).
[31] M. Puri, A. Solanki, T. Padawer, S. M. Tipparaju, W. A. Moreno, Y. Pathak, *Artificial Neural Network For Drug Design, Delivery And Disposition*, Academic Press, Boston **2016**, pp. 3–13.
[32] G. Rizzello, M. A. Mandolino, M. Schmidt, D. Naso, S. Seelecke, An Accurate Dynamic Model for Polycrystalline Shape Memory Alloy Wire Actuators and Sensors, *Smart Mater. Struct.* **2019**, *28*, 025020.
[33] M. A. Mandolino, F. Ferrante, G. Rizzello, A Hybrid Dynamical Model for Hysteretic Thermal Shape Memory Alloy Wire Actuators, *IFAC-PapersOnLine* **2020**, *53*, 8923.
[34] D. E. Rumelhart, G. E. Hinton, R. J. Williams, Learning Internal Representations by Error Propagation, Parallel Distributed Processing, Explorations in the Microstructure of Cognition, *Biometrika* **1986**, *71*, 6.
[35] P. J. Werbos, Backpropagation through Time: What It Does and How to Do It, *Proc. IEEE* **1990**, *78*, 1550.
[36] D. P. Kingma, J. Ba, Adam: A Method for Stochastic Optimization, *International Conference on Learning Representations* **2014**, *abs/1412.6980*, https://api.semanticscholar.org/CorpusID:6628106.
[37] Y. Yu, C. Zhang, X. Zhang, C.-Y. Su, M. Zhou, Iterative Learning Control based on Neural Network and Its Application to Ni-Mn-Ga Alloy Actuator with Local Lipschitz Nonlinearity, *IEEE Trans. Industr. Inform.* **2024**, *20*, 8138.
[38] L. Riccardi, D. Naso, B. Turchiano, H. Janocha, Design of Linear Feedback Controllers for Dynamic Systems with Hysteresis, *IEEE Trans. Control Syst. Technol.* **2013**, *22*, 1268.
[39] D. V. Sabarianand, P. Karthikeyan, Non-Linear Hysteresis Modelling of Piezoelectric Actuator using Feedforward with PI Control for Micromanipulation, *Microsyst. Technol.* **2022**, *28*, 2093.
[40] L. Liu, K. K. Tan, S. Chen, C. S. Teo, T. Heng Lee, Discrete Composite Control of Piezoelectric Actuators for High-Speed and Precision Scanning, *IEEE Trans. Industr. Inform.* **2012**, *9*, 859.

