Distal-end force prediction of tendon-sheath mechanisms for flexible endoscopic surgical robots using deep learning

Li, Xiaoguo; Cao, Lin; Tiong, Anthony Meng Huat; Phan, Phuoc Thien; Phee, Soo Jay

2019

Li, X., Cao, L., Tiong, A. M. H., Phan, P. T., & Phee, S. J. (2019). Distal-end force prediction of tendon-sheath mechanisms for flexible endoscopic surgical robots using deep learning. *Mechanism and Machine Theory*, 134, 323-337. doi:10.1016/j.mechmachtheory.2018.12.035

<https://hdl.handle.net/10356/137858>

<https://doi.org/10.1016/j.mechmachtheory.2018.12.035>

© 2019 Elsevier Ltd. All rights reserved. This paper was published in *Mechanism and Machine Theory* and is made available with permission of Elsevier Ltd.

Downloaded on 26 Apr 2023 20:14:25 SGT

Distal-End Force Prediction of Tendon-Sheath Mechanisms for Flexible Endoscopic Surgical Robots using Deep Learning

Xiaoguo Li, Lin Cao*, Anthony Meng Huat Tiong, Phuoc Thien Phan and Soo Jay Phee

*Robotics Research Center, School of Mechanical & Aerospace Engineering, Nanyang Technological University
50 Nanyang Avenue, Singapore 639798*

*Corresponding to lin.cao@ntu.edu.sg

Abstract

Accurate haptic feedback is highly challenging for flexible endoscopic surgical robots due to space limitation for sensors on small end-effectors and critical force hysteresis of their tendon-sheath mechanisms (TSMs). This paper proposes a deep learning approach to predicting the distal force of TSMs when manipulating a biological tissue based on only proximal-end measurements. Both Multilayer Perceptron (MLP) and Recurrent Neural Network (RNN) were investigated to study their capabilities of making sequential distal force predictions. The results were compared with those of the conventional modelling approach. It was observed that, when sufficient data was provided for training, RNN achieved the most accurate prediction (RMSE=0.0219 N) in experiments with constant system velocity. The effects of insufficient training data, varying system velocity and irregular motion trajectories on the performance of RNN were further studied. Notably, RNN could precisely identify the current system phase in the force hysteresis profile and can be applied to TSMs with realistic non-periodic movement such as manual manipulation trajectory (RSME=0.2287 N). The proposed approach can be applied to any TSM-driven robotic systems for accurate haptic feedback without requiring sensors at the distal ends of the robots.

Keywords—Haptic feedback, Surgical robot, Tendon sheath mechanisms, Deep learning, Hysteresis

1. Introduction

Robotic-assisted Minimally Invasive Surgery (MIS) has been trending in recent years for various surgical procedures such as general surgery, oncological surgery, and colorectal surgery [1]. Robotic-assisted Natural Orifice Transluminal Endoscopic Surgery (NOTES), among all, shows great potential and reliability in performing operations inside the peritoneal cavity while avoiding the necessity of abdominal incisions. Phee et al. [2] developed a master-slave flexible endoscopic surgical robot that greatly improved surgeons' maneuverability within peritoneal cavities. A flexible endoscope is used to guide the slave robot arms to desired surgical sites while the master robot resides at the proximal end to actuate under the command of a surgeon. As the robustness, versatility, and reliability of robot-assisted NOTES rapidly improved over the years, the absence of accurate haptic feedback remained one of the major shortcomings. As a result, surgeons make decisions mainly based on visual feedback and experience [3]. Substantial studies have shown that providing haptic feedback to surgeons will not only reduce the cases of excessive or insufficient force maneuver to diminish tissue damages but also significantly shorten intraoperative time [4-7]. Although there are various commercialized haptic devices available in the market such as Omega.7, CyberForce, and CyberGrasp [5, 8] that can provide decent haptic feedback for surgeons, what is missing from the process is the interaction force information between the surgical robots and operated objects.

In robotic systems for NOTES, Tendon-Sheath Mechanisms (TSMs) have been widely adopted for motion and force transmission because of their high level of flexibility and responsiveness of control in narrow and tortuous paths [2]. However, TSMs are intrinsically associated with issues such as backlash, hysteresis, and nonlinear tension loss due to the friction between the tendon and the associated sheath. As a result, gaining accurate haptic feedback is challenging for these systems. Various sensors, including sensors with working principles based on displacement [9], current [10], pressure [11], resistance [12], capacitance [13], vibration [14] and optical properties [15], may be mounted at the distal-end of a surgical robot to directly measure the interaction force for haptic feedback. However, these sensors are mostly limited by the availability of mounting space at the distal-end, harsh working environment during the surgery, sterilization capability, and issues with the associated wires and accessories, etc.

Alternatively, for TSM-driven robots, extensive efforts have been made to model the force transmission of the robot mathematically so that the force at the distal-end of the robot can be computed through the models using the measurements at the proximal-end. Kaneko et al. [16] analyzed the tension transmission in TSMs based on Coulomb friction model. Inspired by the Coulomb model, Lampaert and Pitkowski et al. [17, 18] have endeavored to modify the modelling technique and proposed Dahl model, LuGre model, and Leuven model. On the other hand, as the fitting performance of these modelling approaches improved, they mainly suffered from discontinuity between different hysteresis phases and inability to accurately describe the friction force when the system operates in the vicinity of zero velocity. In a more recent development of medical robots, Bouc-Wen model [19] has been pervasively used to model the nonlinear friction in TSMs. Such method is often used for backlash compensation in Bowden cable control to reduce hysteresis [20, 21]. Do et al. [22, 23] proposed a modified Bouc-Wen model with dynamic properties that employed velocity and acceleration information to depict the friction profile in an even higher precision manner. It is noted, in the majority of literature, a spring was used to simulate the reaction of tissue. In reality, to practically implement force prediction on a haptic device, the non-linear property of tissue needs to be taken into account. Wang et al. [24] compared tissue modelling methods including Voigt model, Kelvin model, Hunt-Crossley model and proposed a new approach to model the tissue force reaction under a constant indentation speed for TSMs in NOTES. In sum, in order to precisely predict the distal force in TSMs using mathematical models, a set of intricate parameters need to be carefully derived for the equations under the assumptions that the pretension of the tendon is large enough to prevent any system slacking, and the types of interactions with tissue are limited for each viscoelastic model to be built.

To overcome these shortcomings, a data-driven approach is proposed to train a deep Artificial Neural Network (ANN) with sensory data from the proximal-end of a robotic NOTES system. The trained neural network makes dynamic estimations of the distal-end force by reading sensory inputs from the proximal load cell and encoder at each time step. The proposed approach requires neither intricate parameter identification process nor the assumption of no system slacking. Most notably, when sufficient and proper training data was provided, the prediction was not only accurate but also reflected the actual force hysteresis profile of TSMs.

This paper is organized as follows: important features of TSMs and the experimental set-up are introduced in Section 2, followed by the experiment design and detailed approaches explained in Section 3. The validation of the proposed approaches is presented and discussed in Section 4 including experiments on both constant and varying system velocities. Finally, the conclusion and future work are given in Section 5.

2. TSMs and Platform Set-up

2.1. Tendon-Sheath Mechanisms

TSMs usually consist of multiple sets of actuation cables (tendons) and hollow coils (sheath). Fig. 1(a) and (b) show the details of TSMs and how they are integrated with a master-slave NOTES robot [2]. The tendon moves inside the sheath to transmit force from the proximal-end (F_{prox}) to distal-end (F_{dist}). On the other hand, the frictional force between tendon and sheath, defined as $f_{\text{TSM}} = (F_{\text{prox}} - F_{\text{dist}})$, introduces a nonlinear relationship between F_{prox} and F_{dist} . Fig. 1(c) reveals the proximal and distal force hysteresis profile of TSM. There are three distinct phases: ① the pulling phase: the tendon is being stretched, ② the releasing phase: the tendon is being relaxed, and ③ the transition phase: f_{TSM} changes its direction when the system operates near zero velocity, known as the deadzone. Al-Bender et al. [25] stated that the friction f_{TSM} in phase ① and ② depend on the relative displacement between tendon and sheath. Moreover, uniquely, the system enters a pre-sliding regime (micro-displacements) in ③ the transition phase which is characterized by the Stribeck curve [26, 27]. Researchers have found that the friction force f_{TSM} in phase ③ is a function of both displacement and previous history of displacement that form a non-local memory hysteresis [28].

Fig. 1 (a) Schematic of TSM; (b) TSMs in NOTES robot [2]; (c) force hysteresis profile.

Equations (1-4), derived from the mechanics of TSM, were widely used in the literature to model the relationship between F_{prox} and F_{dist} [3, 29, 30]. This modelling approach gains popularity in practice for its ease of in-situ parameter identification and relatively good accuracy, and thus it was employed in this study to demonstrate the accuracy and effectiveness together with the proposed deep-learning approach. In addition, the experiments were conducted under similar assumptions, such as fixed TSM configuration and no sheath buckling. Methods from Refs. [19-23] require complex parameter identification and refinement using ad-hoc combinations of various optimization strategies such as Genetic Algorithm and Nelder-Mead Simplex method. Moreover, the tendon must be maintained by a suitable pretension during the motion as models in Refs. [19-23] work under the assumption of no system slacking [28], whereas, approaches used in this study have no such restriction. This is an important factor to consider as there exists system slacking during the real-life operation of a tendon-driven surgical robot [2]. Therefore, the above models were not selected because the assumptions did not hold.

Equation (1) and (2) compute the distal force $F_{dist,\varphi}$ in the ① pulling and ② releasing phases. Equation (3) shows that the distal force in the ③ transition phase $F_{dist,tran}$ is taken as the distal force of ① pulling phase computed at the last time frame $F_{dist,pull,last}$, denoted in Fig. 1(c). In the transition phase, the proximal force drops but the distal force is assumed constant. Equation (4) computes the corresponding proximal force $F'_{prox,tran}$ at the moment the distal force starts dropping.

$$F_{dist,\varphi} = \begin{cases} F_{prox}e^{-\mu\varphi\xi}, & \text{if } \varphi < \Phi_0 \\ T_0, & \text{if } \varphi \geq \Phi_0 \end{cases} \quad (1)$$

$$\Phi_0 = \min\{\varphi \mid F_{dist,\varphi} = T_0\} \quad (2)$$

$$F_{dist,tran} = F_{dist,pull,last} \quad (3)$$

$$F'_{prox,tran} = F_{prox,tran}e^{-2\mu\varphi} \quad (4)$$

, where φ denotes the accumulated curve angle along the tendon, μ denotes the friction coefficient between the tendon and sheath and ξ denotes the direction of movement. T_0 is the pretension in the system, and $F_{prox,tran}$, $F'_{prox,tran}$ are the initial and final proximal force in the ③ transition phase respectively. Φ_0 represents the maximum length along the tendon until where the input force can be transmitted.

In the experiments of this study, φ was kept at constant 3π by fixing the TSM on the platform as shown in Fig. 2(b). By collecting a set of data on $\langle F_{dist}, F_{prox} \rangle$, the linear regression method was performed on Equation (1) to calculate μ to be 0.07644. $\xi = 1$ in the ① pulling phase and $\xi = -1$ in the ② releasing phases according to the definition. Refer to Ref. [29] for the derivation of the above equations and detailed procedure of parameter identifications.

2.2. Physical Experimental Test Bed

The main purpose of this work was to investigate the performance of the deep ANN approach for TSM distal force estimation during tissue manipulation. As shown in Fig. 2 (b), a test bed was devised to generate training and validation data to examine the performance of deep ANNs experimentally. A DC motor (2657W024C R from Faulhaber Schonaich®) was employed to actuate the tendon from the proximal-end. A motor encoder (HEDS-5540 A14, 500 CPR,

accuracy 0.18°) obtained the actual displacement and velocity of the tendon d_{prox}, v_{prox} . For all experiments, the DC motor was under closed-loop PID displacement control using feedback from the proximal encoder. A 0.45 mm diameter tendon (Asahi Intecc®) was used in the test bed to transmit movement to the distal side via a sheath (Asahi Intecc®) of 1.14mm diameter. Two FUTEK LTH300 (capacity 222 N, accuracy 0.007N) compression load cells were attached to both the proximal and distal ends of the sheath, respectively, to measure the sheath compression force during the operation. It was proven by Wang et al. [30] that the tension in the tendon is equal to the compression force in the sheath. Therefore, F_{prox} and F_{dist} were acquired as the corresponding proximal and distal tension of the tendon. At the distal side, a piece of tissue was mounted between the end of a tendon and a fixed stand. Sensory feedback was acquired and synchronized by MATLAB SIMULINK interfaced with a QPIDE data acquisition board from Quanser Quarc®. We used 1 kHz for sensor reading sampling and analog output processing which is sufficient for the application. All instruments have identical specifications as those used in the actual master-slave endoscopic surgical robot [2] to ensure practicality and repeatability of the experiments. Note that the experiment was conducted with fixed TSM configuration. The accumulated angle of sheath configuration was kept at 3π . Furthermore, we made sure that the sheath does not buckle throughout the experiment so that the data we collect was consistent.

Fig. 2 (a) Schematic of the experiment set-up, (b) experiment test bed for validation

3. Methodology

3.1. Experiment Design

To examine the proposed approach, experiments were designed in two phases. In the first phase, twenty-three pieces of tissue from different parts of two ex vivo porcine stomachs were mounted on the test bed and stretched by the tendon attached. Data $\langle d_{prox}, F_{prox}, F_{dist} \rangle$ was generated while each tissue was pulled by forces of various magnitudes for 200 seconds. The motor trajectory for data generation in the first phase of the experiments is depicted in Fig. 3. The magnitudes of the motor's cycling motion ranged from 5° to 50° with an increment of 2.5° (diameter of the pulley is 26mm), and the tendon was kept at a constant speed (2 mm/s) with closed-loop position control. This speed is normal for tendon-driven endoscopic surgical robots, and the incremental motion magnitude ensured a wide spectrum of proximal motion and force. One set was randomly chosen to be the test data set, two sets were randomly chosen to be the development data sets, the remaining twenty sets were concatenated stochastically with `randperm()` command in MATLAB, and were then used to train the deep neural network models. In this study, we have all taken 200s of testing data, so when comparing result performance, there was no bias introduced from the time duration. However, in practice, the prediction output from the proposed LSTM model can be of any time length. Note that the input of the neural

network $\langle d_{prox}, F_{prox} \rangle$ for training, development and testing sets were normalized with training mean μ_{train} and variance σ_{train}^2 to improve learning efficiency. $\langle F_{dist} \rangle$ was labelled as the target during training. Development sets were used to tune the hyper-parameters of artificial neural networks before we validated on the test set. The trained neural networks made predictions on $\langle \hat{F}_{dist} \rangle$ which were then compared with actual load cell readings to measure their performances.

Fig. 3 Motor movement measured by the proximal encoder.

In the second phase of the experiments, the effects of varying system movement, including velocity and trajectories, were further explored. Force hysteresis profiles under different system speed were plotted to look into the magnitude of their impact. Follow up experiments were carried out to investigate if a neural network trained under constant system velocity was able to make reasonable predictions on the system with varying system velocity and trajectories. Final experiments were designed to train a neural network using non-periodic excitation motor movement with varying velocity $\langle d_{prox}, F_{prox}, v_{prox}, F_{dist} \rangle$. In order to show that the trained model demonstrated good performance not because of overfitting on the training data sets, predictions on $\langle \hat{F}_{dist} \rangle$ were made on both computer generated test data set and manual manipulation test data set generated from random movement of human hands.

3.2. Artificial Neural Networks

The application of machine learning for surgical robots is emerging in recent years [31, 32]. The integration of ANN and surgical robot endow the latter with the capability to gain intuition from processing vast amounts of data [33] and learning from human demonstration [34], both of which are extremely difficult to implement by hard programming the robots [35]. Two suitable types of ANN are considered for our approach.

Fig. 4 Proposed scheme for MLP prediction.

3.2.1. Multilayer Perceptron

Multilayer Perceptron (MLP) is a class of ANN commonly used as a supervised learning technique for classification and regression problems [36]. MLP is a popular choice when dealing with forecasting problems for its outstanding ability to learn and generalize the non-linear relationship between input and target sets.

The information between the neurons is computed as follows:

$$output_j^{(n)} = \sigma^{(n)}(\sum_i output_i^{(n-1)} W_{ij}^n + b_{ij}^n) \quad (4)$$

The $output_j^{(n)}$ is defined as the output of $neuron_j$ in $layer^{(n)}$ computed from the activation function $\sigma^{(n)}(\cdot)$. $W_{ij}^{(n)}$, $b_{ij}^{(n)}$ are the weight matrix and bias associated with $neuron_j$ and $neuron_i$ in the previous layer.

Filtered load cell and encoder readings were fed into the input layer after normalization for each time step. Equation (4) was carried out whenever information was passed from one neuron to the connected neurons in the next layer. As illustrated in Fig. 4, an MLP was designed to make predictions on the distal force $\langle \hat{F}_{dist} \rangle$ after training with $\langle d_{prox}, F_{prox}, F_{dist} \rangle$ from 20 data sets. The choice of the activation function $\sigma^{(n)}(\cdot)$ and other hyper-parameters that we set for the network will be presented in Section 4.

Fig. 5 Proposed scheme for LSTM prediction.

3.2.2. Recurrent Neural Network

Similar to speech recognition and language translation, the input and output signals in the problem we sought to solve were time-series sequential data. Long Short-Term Memory (LSTM) is a type of recurrent neural network that stands out when tackling time series problems [37, 38]. In traditional feedforward neural networks, like convolutional neural networks and MLP, information only flows in one direction, from input to output. However, in a recurrent net, the output of a layer is fed back into the same layer together with inputs of the next time step. With the built-in feedback loop, for each time step, LSTM takes not only inputs from the current time step but also the information from all previous time steps, allowing it to act as a forecasting engine with a memory mechanism. In addition, LSTM has a set of forget gate, input gate and output gate to regulate information going in and out of the cell state $C^{(t)}$ to handle the long-term dependencies [39].

Fig. 5 demonstrates the architecture of LSTM used in this work when unfolding in time. The LSTM takes input signals of the current time step from the load cell and encoder along with the previous prediction $\hat{F}_{dist}^{(t-1)}$ to make a prediction $\hat{F}_{dist}^{(t)}$ for the current time step. The forget gate, input gate, and output gate are denoted by 'f', 'i' and 'o' respectively. Each of these gates has their weight matrix and bias.

The forget gate learns how much to forget from previous cell state and the input gate determines how much new information to let in for a current time step, where W_{cand} and b_{cand} are weight and bias for candidate layer tanh:

$$f^{(t)} = \text{sigmoid}(W_f \cdot \langle \hat{F}_{dist}^{(t-1)}, d_{prox}^{(t)}, F_{prox}^{(t)} \rangle + b_f) \quad (5)$$

$$i^{(t)} = \text{sigmoid}(W_i \cdot \langle \hat{F}_{dist}^{(t-1)}, d_{prox}^{(t)}, F_{prox}^{(t)} \rangle + b_i) * \tanh(W_{cand} \cdot \langle \hat{F}_{dist}^{(t-1)}, d_{prox}^{(t)}, F_{prox}^{(t)} \rangle + b_{cand}) \quad (6)$$

The cell state is the long-term memory; it represents all of the learning done in previous time steps:

$$C^{(t)} = f^{(t)} * C^{(t-1)} + i^{(t)} \quad (7)$$

Finally, the output gate is the attention mechanism that decides how much information to output from the cell state $C^{(t)}$ and $\hat{F}_{dist}^{(t)}$ is computed through an identity activation function $g(\cdot)$:

$$o^{(t)} = \text{sigmoid}(W_o \cdot \langle \hat{F}_{dist}^{(t-1)}, d_{prox}^{(t)}, F_{prox}^{(t)} \rangle + b_o) * \tanh(C^{(t)}) \quad (8)$$

$$\hat{F}_{dist}^{(t)} = g(W_F \cdot o^{(t)} + b_F) \quad (9)$$

The prediction $\hat{F}_{dist} = \langle \hat{F}_{dist}^{(1)}, \hat{F}_{dist}^{(2)}, \dots \dots \rangle$ forms a smooth signal to reflect algorithm's prediction at each time step.

4. Results and Discussion

4.1. Phase I

As mentioned previously, the tests aimed to simulate and compare the performance of two types of deep neural networks regarding their prediction accuracy and abilities to generalize the hysteresis profile of TSMs. In this section, MLP and RNN-LSTM were examined experimentally using identical training, testing and development data sets. Together their results were compared with those of the mathematical modelling approach conventionally used in the literature. After that, the better ANN was further investigated in its repercussion provoked by using a lessened training data set and the possibility of applying transfer learning.

4.1.1. Performance of LSTM, MLP, and Modelling

	Number of hidden layers	Number of neurons in each layer	Activation function	RMSE
LSTM	4	65	Tanh	0.0219 N
MLP	4	100	ReLU	0.1160 N
Modelling	N.A			0.4729 N

Table 1 Architecture and performance.

Both MLP and RNN-LSTM were built in deep models and their final architectures used in this paper are listed in Table 1. Development sets were used in this case to tune the hyper-parameters of ANNs so that the information in the test set does not infiltrate into the tuning process. Since finding the optimal architectures for ANNs is still an open research, the authors followed the rule of thumb to choose the hyper-parameters [40, 41]. The following range of tuning is recommended for this particular problem scale: {2-4} and {50-150} for hidden layers and neurons in each layer, respectively. ANNs were trained with 0.001 learning rate for 1000 epochs using Adam optimizer [42]. All weights in hidden layers were randomly initialized using Glorot uniform distribution. In addition, 5% dropout was applied to both ANNs during training to prevent overfitting.

RMSE was calculated using $\langle F_{dist} \rangle$ from distal load cell readings and $\langle \hat{F}_{dist} \rangle$ produced by corresponding methods over the 200s test data set. Both ANNs were found to produce good tracking and smaller RMSE than using the mathematical modelling approach. For a clear view, only three cycles were plotted in Fig. 6(a) to observe how well the prediction signals reflect the distal force of TSM. All three methods were able to depict the overall pattern, yet it has to be highlighted that LSTM was in best agreement with the actual load cell readings. In addition, as shown in Fig. 6(b) hysteresis profile, LSTM and the modelling approach captured more precise detail of the transition phase than MLP. It has come to our attention that since the modelling approach applied different equations for different hysteresis phases there were discontinuities in between transition phases and releasing phases as shown in the zoom-in view of Fig. 6(d). In comparison, hysteresis profiles learnt by two ANNs approaches demonstrated full continuity throughout the entire prediction process. Our experiment results evidentially proved that LSTM performed outstandingly in both prediction accuracy and identifying of the actual hysteresis profile.

The transition phase is crucial to the force prediction process as it usually contains the local maximal force. It was clear that both ANNs made excellent predictions in the pulling and releasing phases. However, the discrepancy in their RMSE mainly came from how the two ANNs handle the transition phase. In Fig. 6(c) zoom-in plot, MLP prediction signal oscillated and tended to overpredict in the transition phase. This result agreed with their learnt hysteresis profile that we observed in Fig. 6(b). The prediction of MLP in this very phase failed to generalize to a consistent pattern, thus, further strengthened the above verdict of the inability of MLP to handle the transition phase signal.

Our study provided insight into how the fundamental difference between MLP and LSTM affected their performance when dealing with time-series sequential data. Most importantly, when a prediction was made on one time step, MLP merely considered inputs of the current time step whereas LSTM's recurrent structure allowed it to make use of not only the information of the current time step but also the information in all previous time steps. The cell state $C^{(t)}$ contains a long-term memory of how the pattern of data would change in accordance with the current input and previous outputs. LSTM firstly identified the hysteresis phase at the present time step, and then with the help of 3 gates regulating information flow accordingly, a good prediction was made that fulfilled the actual force hysteresis profile. It was difficult for MLP to do the same because by considering the inputs of the only current time step, there was not enough information to realistically identify the current force hysteresis phase. As a result, a feedforward net, like MLP, would process the input sensory data from different hysteresis phases in the same manner. Moreover, as aforementioned in Section 2.1 Tendon-Sheath Mechanism, the friction loss between the F_{prox} and F_{dist} during transition phase depends on the previous history of displacement. Therefore, the memory mechanism of LSTM was critical in this case to learn and generalize the distal force behaviors during the transition phase.

Fig. 6(a) Prediction results (for a clear view, only three cycles are plotted); (b) predicted hysteresis profiles; (c) zoom-in of one ANN prediction peak; (d) zoom-in of modelling approach discontinuity between transition phase and releasing phase.

4.1.2. Further Investigation in Choice of Training Data

Our results provided compelling evidence that the proposed LSTM approach is fully competent of learning the force hysteresis model of TSMs with sufficient training data. On the other hand, acquiring training data from 20 pieces of tissue for LSTM was feasible but time-consuming. A further experiment was carried out to test the ability of LSTM's to learn based on data generated from only one tissue and generalize on other tissue.

Fig. 7 presents the results of LSTM trained with data generated from one random tissue and predicted on the benchmark tissue (same test tissue used in Section 4.1.1). Prediction RMSE was 0.1072 N. Identical experiments were carried out on three other pairs of different tissue. Similar prediction performance was obtained with an average RMSE of 0.1364 N. It was clear that LSTM still managed to produce a decent prediction accuracy using the lessened training data set while learning the correct force hysteresis profile in different phases.

We observed that if an LSTM model trained with the data generated from one tissue were used to make predictions on the same training tissue, i.e., the training data set was identical to the testing data set. The average prediction performance was $RMSE = 0.0539$ N. The prediction accuracy was distinctly good. However, it wasn't as good as the LSTM trained with 20 different tissue ($RMSE = 0.0219$ N). Besides, unlike the former, training data set for 20-tissue-LSTM model didn't contain the testing data set which makes it a more difficult problem setup.

To sum up, training an LSTM model using a diversified training data set is crucial for making accurate predictions. Nevertheless, LSTM is capable of learning the actual force hysteresis profile from lesser training data and make a moderately good prediction.

Fig. 7(a) LSTM prediction; (b) hysteresis profile.

4.1.3. Transfer Learning

Transfer learning is a research problem in machine learning that focuses on storing knowledge learned from one problem (source task) and applying it to a different but related problem (target task). Such an approach brings benefits such as avoiding extensive data acquiring efforts [43].

It would make sense to apply transfer learning if the source task (ST) and target task (TT) fulfill the following assumptions:

- ST and TT share the same inputs. In this case, inputs are the proximal load cell and encoder signals.
- Low-level features from ST are helpful for learning TT. In our case, how the distal force changes in accordance with the input signals.
- Training data for ST are more accessible than training data for TT.

Fig. 8(a) Transfer learning prediction; (b) hysteresis profile.

To explore the possibility of applying transfer learning, another effort was made to train an LSTM model using data generated from manipulating a spring at the distal-end and then make predictions on the benchmark tissue (same test tissue used in Section 4.1.1). The stiffness of the spring and benchmark tissue was measured as 18.89 N/m and 4.80 N/m, respectively. Note that the stiffness of spring and tissue was not similar. However, this does not prevent LSTM from learning the relationship between the proximal and distal force in a given TSM.

Prediction results showed that the model learnt from manipulating a spring (ST) can be successfully transferred to make distal force predictions on tissue manipulation (TT). Fig. 8 shows that the overall pattern (low-level feature) was well predicted. However, RMSE of 0.1838 N was measured, and there was obvious error in the rear pulling and releasing phase. Namely, the response when the payload tension was high and when the payload was relaxing back to its non-tensioned form. In both cases, tissue's viscoelasticity was proved to bring a difference to the high-level features of force response.

We conclude that LSTM could make satisfactory predictions on tissue force response by learning from force response of a spring. However, it was difficult to eliminate the prediction error completely due to the intrinsic disparity between the tissue and spring's viscoelastic property. The force response difference caused by such disparity was easily reflected on the distal load cell, as high-level features, but not so easily on the proximal load cell because of the friction loss in TSM. As a result, during transfer learning, LSTM was only trained with the high-level features of spring force response but not from those of the tissue. Note that the current accuracy of transfer learning does not surpass the accuracy of the mathematical modelling approach. However, the authors believe that this could be a promising direction for further study so that we can achieve a short data collection procedure and good prediction accuracy. Possible approaches to tackle the problem with lacking of high-level features are discussed in Section 5 future work.

4.2. Phase II

In this part of the research, the necessity of considering the impact of training tendon movement, including velocity and trajectories, was further investigated. First of all, an experiment was conducted to examine force hysteresis profiles of the same TSM setup operating under different system speeds.

The normal system velocity of the tendon-driven NOTES robot ranges from 0.5 mm/s to 10 mm/s. Fig. 9 reveals the five force hysteresis profiles of a TSM operating under an identical tendon displacement but different system operation speed. It was observed that the increment of velocity brought two distinct effects to the hysteresis profile: 1. slower slope in the pulling phase and steeper slope in the releasing phase; 2. decrease of the distal force in the transition phase. It was also observed that the bigger difference in velocity the more obvious the above effects occurs. It was clear that varying velocity affects the force hysteresis profiles of TSM yet the magnitude of impact on the performance of the deep LSTM model remains unknown.

Therefore, another experiment was designed to study how much would the varying velocity affects the accuracy of deep LSTM in distal force prediction tasks.

Fig. 9 Impact of velocity on force hysteresis profiles.

4.2.1. The Impact of Varying System Velocity

To test the impact of varying velocity, an LSTM was trained using data sets generated under system velocity of 2mm/s. The trained LSTM model was then used to make predictions on 5 different test data sets with velocities of 0.5mm/s, 1mm/s, 2mm/s, 5mm/s, and 10mm/s, individually. Note that the tendon movements for all data sets collected shared identical trajectories. The only difference lies in their system velocities. In addition, a spring (stiffness 18.89 N/m) was used for experiments in Section 4.2.1 and 4.2.2 as payload at the distal end. The goal is to control the variables so that the response of the payload does not introduce bias.

Fig. 10 plots the force hysteresis profile prediction results of the LSTM on test sets of varying velocities. Fig. 10(c) shows good prediction accuracy since the test velocity is the same as the training velocity. Evidentially, smaller system velocities, shown in Fig. 10(a-b), demonstrated similar prediction patterns and so did the ones with larger system velocities, shown in Fig. 10(d-e). Furthermore, as we compared Fig. 10(a-b) to Fig. 10(d-e), it could be observed that increased and decreased system velocities resulted in different types of error patterns. Error magnitude increases with the difference of velocity between training and testing. RMSEs for data sets plotted in Fig.10 (a-e) are 0.4552 N, 0.3454 N, 0.0469 N, 0.5704N and 1.4920 N, respectively. The influence of velocity on LSTM's prediction accuracy was clear.

Fig. 10 Prediction results for test sets with different velocities (the LSTM model was trained with data sets of 2mm/s)

4.2.2. The Motion Trajectories

Inspired by experiments in 4.2.1, we also noticed that the trajectories of the tendon movement could also be a variable that affects LSTM's learning. To test the impact of varying trajectories, an LSTM was trained using data sets generated from constant speed (2 mm/s) motor trajectories, shown in Fig. 11(a), and predicted on test data set generated from sine wave motor trajectories, shown in Fig. 11(b). The experiment was designed in a way that the sine wave trajectories speed is close to 2 mm/s when pulling and releasing.

As shown in Fig. 11(c), for LSTM prediction, it was observed that there was over-prediction in the higher peaks (average peak error magnitude of 40.7%) and under-prediction in the lower peaks (average peak error magnitude of

39.4%). During the range of peaks, the difference in trajectories is bigger, and the trained LSTM made more significant prediction error than it would during the range of pulling and releasing phase. The mathematical modelling approach was also applied to the test data set and its prediction is shown together in Fig. 11(c). The RMSEs for LSTM and modelling approach predictions were 0.7512 N and 0.6136 N, respectively. Clearly, the trajectories in Fig. 11(a) needed to be improved so as to train a more versatile LSTM model.

The experiment results in 4.2.1 and 4.2.2 have surfaced the incompleteness of training data generation if the impact of system velocity and trajectories were not taken into account. In addition, we concluded that it was necessary to consider the varying velocity when designing training data generation for LSTM and we should use $\langle d_{prox}, F_{prox}, v_{prox} \rangle$ as input instead of $\langle d_{prox}, F_{prox} \rangle$.

Fig. 11(a) Training data: motor movement measured by the proximal encoder; (b) testing data: motor movement measured by the proximal encoder; (c) LSTM prediction.

Fig. 12(a) Motor movement; (b) actual proximal and distal force.

Fig. 13(a) LSTM prediction on non-periodic movement; (b) hysteresis profile.

Fig. 14 LSTM prediction on manual manipulation.

4.2.3. Training with Non-periodic Motor Movement

In order to train an LSTM model to make good predictions on a system with varying velocity and trajectories, a more versatile training strategy was proposed to incorporate the computer-generated, non-periodic motor movement, shown in Fig. 12(a). The input of LSTM was set as $\langle d_{prox}, F_{prox}, v_{prox} \rangle$ so that the neural network can learn and generalize the relationship between system velocity and force hysteresis profile during training. As presented in Fig. 12(b) and Fig. 13(b), the problem setup became harder as there were occasions when the proximal load cell perceived significant fluctuations in force yet distal force remained almost constant. The force hysteresis profile no longer follow the pulling-transition-releasing pattern in sequence, and can sometimes reverse back to pulling phase from the transition phase. Therefore, identifying the correct hysteresis phase is paramount. This phenomenon was mainly caused by the irregular movement of the TSM and occurred mostly during the prolonged transition phase region.

To examine the proposed strategy, twenty-three pieces of tissue were used to generate experimental data. Three of them were randomly chosen to be the testing data sets, and the remaining twenty pieces were used as the training data sets. Fig. 13 demonstrates the prediction performance and hysteresis profile learnt by LSTM. The best prediction accuracy among the three testing data sets has an RMSE of 0.0917 N, and the average RMSE is 0.0975 N. Note that the overall prediction was in good agreement with the actual load cell readings yet not as good as the prediction done on the constant velocity experiment set-up. Specifically, the biggest prediction error was mainly introduced from two regions:

- Higher peaks: max error occurred when error magnitude = 0.42 N and error = 3.2%
- Lower peaks: max error occurred when error magnitude = 0.21 N and error = 4.5%

From Fig. 13(b) hysteresis profile, we found that LSTM was capable of adapting to different pulling phase and releasing phase slope angles when the TSM was operating under various system velocities. In addition, the critical points that the system switches into transition phases were accurately located. However, small discrepancies were found in the prediction for the transition phase.

To sum up, LSTM gave good predictions on identifying the correct hysteresis phase that the system was in and the overall performance was improved when taking into account of the velocity during the training data collection.

4.2.4. Predictions on Manual Manipulation

During the actual operation of robotic-assisted NOTES, the motor movement is not as smooth as computer-generated trajectories. The irregular movement trajectories and the jerkiness of human hand manipulation must be considered. The last experiment in this study was done to examine how well the model trained in Section 4.2.3 would perform on a more realistic case when tissue was manipulated under a manually controlled motor movement through TSM. The manual control signal is generated by operating an Omega.7 haptic device (Force Dimension, Switzerland), mimicking the motion trajectory and velocity in real surgical operations. Prediction performance was depicted in Fig. 14. It was clear that we can apply the LSTM model trained with computer generated training data sets and make reasonable predictions on human hand manipulation on the real tissue. RMSE of the prediction was 0.2287 N with the maximum error of 0.81 N.

In this study, manual manipulation data was not considered as training data sets for LSTM. This is because the data generation process could be time-consuming and tedious for end users to conduct. The strategy proposed in Section 4.2.2 could produce decent prediction accuracy, so it would be the preferred approach for its simplicity.

The time need for data collection on twenty tissue during real-life use is around 3 hours. Training an LSTM modelling based on the above training data using the architecture described in this paper requires 4 to 5 hours. For testing, the average time taken to make one prediction is 0.00075s for a single time step. This information was measured with our validation experimental set-up. Our PC configuration was as follows: Processor-Intel(R) Xeon(R) CPU E5-1650 v3 @ 3.5GHz, RAM-32.0 GB, GPU-NVIDIA Quadro K2200.

5. Conclusion

This study investigated the performance of deep ANNs on predicting the distal-end force of TSMs based on only proximal-end sensory data from surgical robotics system. When sufficient data was provided for the training, the proposed RNN-LSTM approach produced good force prediction results with RMSE of 0.0219 N and 0.0975 N on the system with a constant operation velocity and varying operation velocity, respectively. LSTM demonstrated superior prediction accuracy during the system transition phases and better capability in identifying the actual force hysteresis profile of TSM when compared with the performance of the mathematical modelling approach and MLP. In addition, applying deep learning to force prediction enjoys the benefit of avoiding complicated parameter identification process and manually specifying the transition phases. It was also perceived that training LSTM with deficient or improper data would distinctively degrade its performance. By incorporating non-periodic motion trajectory with varying velocity in training data collection, the accuracy of the prediction for more realistic manipulations was significantly improved. Moreover, to reduce the training effort, it was possible to train an LSTM model with force response of a spring and make a satisfactory prediction for manipulations on biological tissue.

The development of deep ANN approach reinforces surgeons' surgical capabilities and shortens the intraoperative time by adding force feedback to haptic devices while avoiding integrating additional sensors at the distal ends of surgical robots. For future work, we would explore the possibility of improving transfer learning technique by enabling the LSTM to learn the low-level features from spring training and then learn the high-level features from a small amount of tissue training. More experiments would be performed to help develop a method to close the prediction gap in the transition phase when the system is operating with changing velocity. Another promising direction of future work would be to make distal force forecast ahead of time. Such an application is possible with the help of action indicators, such as modeled surges, which can be extracted from the surgeons' input by using Gaussian Mixture Models [32]. Moreover, further studies are to be carried out to further extend the proposed approach for TSMs with varying sheath configurations.

Acknowledgments

Research supported by the National Research Foundation (NRF) Singapore (NRFI2016-07).

References

- [1] G. G. Hamad and M. Curet, "Minimally invasive surgery," *The American Journal of Surgery*, vol. 199, no. 2, pp. 263-265, 2010.
- [2] S. J. Phee, S. C. Low, V. A. Huynh, A. P. Kencana, Z. L. Sun, and K. Yang, "Master and slave transluminal endoscopic robot (MASTER) for natural Orifice Transluminal Endoscopic Surgery (NOTES)," in 2009 Annual International Conference of the IEEE Engineering in Medicine and Biology Society, 2009, pp. 1192-1195.
- [3] Z. Wang, Z. Sun, and S. J. Phee, "Haptic feedback and control of a flexible surgical endoscopic robot," *Computer Methods and Programs in Biomedicine*, vol. 112, no. 2, pp. 260-271, 2013/11/01/ 2013.
- [4] S. Ehrampoosh, M. Dave, M. A. Kia, C. Rablau, and M. H. Zadeh, "Providing haptic feedback in robot-assisted minimally invasive surgery: A direct optical force-sensing solution for haptic rendering of deformable bodies," *Computer Aided Surgery*, vol. 18, no. 5-6, pp. 129-141, 2013/09/01 2013.
- [5] A. M. Okamura, "Haptic Feedback in Robot-Assisted Minimally Invasive Surgery," *Current opinion in urology*, vol. 19, no. 1, pp. 102-107, 2009.
- [6] T. Akinbiyi et al., "Dynamic Augmented Reality for Sensory Substitution in Robot-Assisted Surgical Systems," in 2006 International Conference of the IEEE Engineering in Medicine and Biology Society, 2006, pp. 567-570.
- [7] M. Tavakoli, A. Aziminejad, R. V. Patel, and M. Moallem, "Methods and mechanisms for contact feedback in a robot-assisted minimally invasive environment," *Surgical Endoscopy And Other Interventional Techniques*, vol. 20, no. 10, pp. 1570-1579, 2006/10/01 2006.
- [8] H. Vincent, R. A. Oliver, C. H. Manuel, G. Danny, and R. D. L. T. Gabriel, "Haptic interfaces and devices," *Sensor Review*, vol. 24, no. 1, pp. 16-29, 2004/03/01 2004.
- [9] J. Rosen, B. Hannaford, M. P. MacFarlane, and M. N. Sinanan, "Force controlled and teleoperated endoscopic grasper for minimally invasive surgery - Experimental performance evaluation," (in English), *IEEE Transactions on Biomedical Engineering*, Article vol. 46, no. 10, pp. 1212-1221, 1999.
- [10] G. Tholey, A. Pillarisetti, W. Green, and J. P. Desai, "Design, Development, and Testing of an Automated Laparoscopic Grasper with 3-D Force Measurement Capability," *Berlin, Heidelberg*, 2004, pp. 38-48: Springer Berlin Heidelberg.
- [11] K. Tadano and K. Kawashima, "Development of a master-slave system with force-sensing abilities using pneumatic actuators for laparoscopic surgery," (in English), *Advanced Robotics*, Article vol. 24, no. 12, pp. 1763-1783, 2010.
- [12] P. Valdastri et al., "Integration of a Miniaturised Triaxial Force Sensor in a Minimally Invasive Surgical Tool," *IEEE Transactions on Biomedical Engineering*, vol. 53, no. 11, pp. 2397-2400, 2006.
- [13] R. D. Howe, W. J. Peine, D. A. Kantarinis, and J. S. Son, "Remote palpation technology," *IEEE Engineering in Medicine and Biology Magazine*, vol. 14, no. 3, pp. 318-323, 1995.
- [14] T. Ohtsuka, A. Furuse, T. Kohno, J. Nakajima, K. Yagyu, and S. Omata, "Application of a new tactile sensor to thoracoscopic surgery: Experimental and clinical study," *The Annals of Thoracic Surgery*, vol. 60, no. 3, pp. 610-614, 1995/09/01/ 1995.
- [15] W. Lai, L. Cao, Z. Xu, P. T. Phan, P. Shum, and S. J. L. Phee, "Distal End Force Sensing with Optical Fiber Bragg Gratings for Tendon-Sheath Mechanisms in Flexible Endoscopic Robots," in *IEEE International Conference on Robotics and Automation (ICRA)*, Brisbane, Australia, 2018.
- [16] M. Kaneko, M. Wada, H. Maekawa, and K. Tanie, "A new consideration on tendon-tension control system of robot hands," in *Proceedings of the 1991 IEEE International Conference on Robotics and Automation*, Piscataway, NJ, United States Sacramento, CA, USA, 1991, vol. 2, pp. 1028-1033: Publ by IEEE.
- [17] V. Lampart, J. Swevers, and F. Al-Bender, "Modification of the Leuven integrated friction model structure," *IEEE Transactions on Automatic Control*, vol. 47, no. 4, pp. 683-687, 2002.
- [18] T. Piatkowski, "Dahl and LuGre dynamic friction models — The analysis of selected properties," *Mechanism and Machine Theory*, vol. 73, pp. 91-100, 2014/03/01/ 2014.
- [19] T. N. Do, T. Tjahjowidodo, M. W. S. Lau, and S. J. Phee, "Nonlinear friction modelling and compensation control of hysteresis phenomena for a pair of tendon-sheath actuated surgical robots," (in English), *Mechanical Systems and Signal Processing*, vol. 60-61, pp. 770-784, Aug 2015.
- [20] B. K. Dinh, L. Cappello, M. Xiloyannis, and L. Masia, "Position control using adaptive backlash compensation for bowden cable transmission in soft wearable exoskeleton," in *2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2016, pp. 5670-5676.
- [21] T. N. Do, T. Tjahjowidodo, M. W. S. Lau, and S. J. Phee, "An investigation of friction-based tendon sheath model appropriate for control purposes," *Mechanical Systems and Signal Processing*, vol. 42, no. 1, pp. 97-114, 2014/01/01/ 2014.
- [22] T. N. Do, T. Tjahjowidodo, M. W. S. Lau, and S. J. Phee, "A new approach of friction model for tendon-sheath actuated surgical systems: Nonlinear modelling and parameter identification," *Mechanism and Machine Theory*, vol. 85, pp. 14-24, 2015/03/01/ 2015.
- [23] T. N. Do, T. Tjahjowidodo, M. W. S. Lau, T. Yamamoto, and S. J. Phee, "Hysteresis modeling and position control of tendon-sheath mechanism in flexible endoscopic systems," *Mechatronics*, vol. 24, no. 1, pp. 12-22, 2014/02/01/ 2014.
- [24] Z. Wang, Z. L. Sun, and S. J. Phee, "Haptic feedback and control of a flexible surgical endoscopic robot," (in English), *Computer Methods and Programs in Biomedicine*, vol. 112, no. 2, pp. 260-271, Nov 2013.
- [25] F. Al-Bender, W. Symens, J. Swevers, and H. Van Brussel, "Theoretical analysis of the dynamic behavior of hysteresis elements in mechanical systems," *International Journal of Non-Linear Mechanics*, vol. 39, no. 10, pp. 1721-1735, 2004/12/01/ 2004.
- [26] B. Armstrong-Hélouvry, P. Dupont, and C. C. De Wit, "A survey of models, analysis tools and compensation methods for the control of machines with friction," *Automatica*, vol. 30, no. 7, pp. 1083-1138, 1994/07/01/ 1994.
- [27] F. Al-Bender and W. Symens, "Characterization of frictional hysteresis in ball-bearing guideways," *Wear*, vol. 258, no. 11, pp. 1630-1642, 2005/06/01/ 2005.
- [28] D. D. T. Nho, "Dynamic Friction Modeling and Control for Flexible Tendon-Sheath Mechanism," Ph.D. Thesis, Robotics Research Centre, School of Mechanical and Aerospace Engineering, Nanyang Technological University, 2014.
- [29] S. J. Phee, S. C. Low, P. Dario, and A. Menciassi, "Tendon sheath analysis for estimation of distal end force and elongation for sensorless distal end," (in English), *Robotica*, vol. 28, pp. 1073-1082, Dec 2010.
- [30] Z. L. Sun, Z. Wang, and S. J. Phee, "Elongation Modeling and Compensation for the Flexible Tendon-Sheath System," (in English), *Ieee-Asme Transactions on Mechatronics*, vol. 19, no. 4, pp. 1243-1250, Aug 2014.

- [31] H. Mayer, F. Gomez, D. Wierstra, I. Nagy, A. Knoll, and J. Schmidhuber, "A System for Robotic Heart Surgery that Learns to Tie Knots Using Recurrent Neural Networks," in 2006 IEEE/RSJ International Conference on Intelligent Robots and Systems, 2006, pp. 543-548.
- [32] C. E. Reiley, E. Plaku, and G. D. Hager, "Motion generation of robotic surgical tasks: Learning from expert demonstrations," in 2010 Annual International Conference of the IEEE Engineering in Medicine and Biology, 2010, pp. 967-970.
- [33] W. Xu, J. Chen, H. Y. K. Lau, and H. Ren, "Data-driven methods towards learning the highly nonlinear inverse kinematics of tendon-driven surgical manipulators," *The International Journal of Medical Robotics and Computer Assisted Surgery*, pp. n/a-n/a, 2016.
- [34] S. Calinon, D. Bruno, M. S. Malekzadeh, T. Nanayakkara, and D. G. Caldwell, "Human-robot skills transfer interfaces for a flexible surgical robot," *Computer Methods and Programs in Biomedicine*, vol. 116, no. 2, pp. 81-96, 2014/09/01/ 2014.
- [35] Y. Kassahun et al., "Surgical robotics beyond enhanced dexterity instrumentation: a survey of machine learning techniques and their role in intelligent and autonomous surgical actions," *International Journal of Computer Assisted Radiology and Surgery*, vol. 11, no. 4, pp. 553-568, 2016/04/01 2016.
- [36] F. Murtagh, "Multilayer perceptrons for classification and regression," *Neurocomputing*, vol. 2, no. 5, pp. 183-197, 1991/07/01/ 1991.
- [37] J. Connor and L. Atlas, "Recurrent neural networks and time series prediction," in *IJCNN-91-Seattle International Joint Conference on Neural Networks*, 1991, vol. i, pp. 301-306 vol.1.
- [38] F. A. Gers, J. Schmidhuber, and F. Cummins, "Learning to forget: Continual prediction with LSTM," 1999.
- [39] S. Hochreiter, Y. Bengio, P. Frasconi, and J. Schmidhuber, "Gradient flow in recurrent nets: the difficulty of learning long-term dependencies," 2001.
- [40] R. Jozefowicz, W. Zaremba, and I. Sutskever, "An empirical exploration of recurrent network architectures," in *International Conference on Machine Learning*, 2015, pp. 2342-2350.
- [41] J. Bergstra and Y. Bengio, "Random search for hyper-parameter optimization," *Journal of Machine Learning Research*, vol. 13, no. Feb, pp. 281-305, 2012.
- [42] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," *arXiv preprint arXiv:1412.6980*, 2014.
- [43] S. J. Pan and Q. Yang, "A Survey on Transfer Learning," *IEEE Transactions on Knowledge and Data Engineering*, vol. 22, no. 10, pp. 1345-1359, 2010.