

Registration of Deformed Tissue: A GNN-VAE Approach with Data Assimilation for Sim-to-Real Transfer

Mehrnoosh Afshar¹, Tyler Meyer², Ron Sloboda³, Siraj Husain², Nawaid Usmani³, and Mahdi Tavakoli¹, *Senior Member, IEEE*

Abstract—In image-guided surgery, deformation of soft tissues can cause substantial errors in targeting internal targets, since deformation can affect the translation of preoperative image-based surgical plans during surgery. Having a realistic tissue deformation simulator could enhance the accuracy of internal targets localization by giving an accurate estimation of the deformation applied to a preoperative model of the organ. A key challenge is to address the sim-to-real gap between the simulator and the actual intraoperative behaviour of the tissue. The sim-to-real transfer challenge is addressed by formulating the problem as a probabilistic inference over a low-dimensional representation of deformed objects. The proposed method utilizes a generative variational autoencoder structure based on graph neural networks (GNN-VAE) to generate a probabilistic low-dimensional representation of the outputs of a physics-based simulator. To match simulation data to real data, the resultant low-dimensional distribution (i.e., prior distribution) is updated iteratively using an Ensemble Smoother with Multiple Data Assimilation (ES-MDA). The advantages of the proposed method are 1) it only uses simulation data for training the GNN-VAE, and no retraining of GNN-VAE is required intraoperatively, 2) it does not require estimating the mechanical properties of the tissue it is simulating, and 3) is able to work with any physics-based simulator. The proposed framework was verified both in experimental and simulation studies and showed it can reduce the registration error in tissue deformation.

I. INTRODUCTION

As part of surgical procedures and interventions such as resection, ablation, biopsy, and brachytherapy, it is important to track the intraoperative positions of lesions and other relevant internal structures. For instance, low-dose-rate, permanent-seed (LDR-PS) breast brachytherapy (a representative surgery in this paper) is a form of cancer treatment that involves the deposition of radioactive seeds into or around tumours or seromas (hollowed-out tumours). During breast brachytherapy,

This research was supported by the Canada Foundation for Innovation (CFI), the Natural Sciences and Engineering Research Council (NSERC) of Canada, the Canadian Institutes of Health Research (CIHR), and the Alberta Jobs, Economy and Innovation Ministry's Major Initiatives Fund to the Center for Autonomous Systems in Strengthening Future Communities.

¹Mehrnoosh Afshar and Mahdi Tavakoli are with the Department of Electrical and Computer Engineering, University of Alberta, AB, Canada T6G 1H9. afsharbo@ualberta.ca, mahdi.tavakoli@ualberta.ca

²Tyler Meyer and Siraj Husain are with the Division of Radiation Oncology, Tom Baker Cancer Centre, 331 29th Street NW, Calgary, Alberta T2N 4N2. tyler.meyer@albertahealthservices.ca, siraj.husain@albertahealthservices.ca.

³Ron Sloboda and Nawaid Usmani are with the Department of Oncology, Cross Cancer Institute, 11560 University Avenue, Edmonton, AB, Canada, T6G 1Z2. nawaid.usmani@albertahealthservices.ca, ron.sloboda@albertahealthservices.ca.

displacements of up to 7 mm are common in the target area [1], [2]. Off-target implantation of radioactive seeds results in insufficient radiotherapy and an increased risk of cancer recurrence. Ultrasound imaging (US) is commonly used to identify and track lesions and targets during interventional procedures; however, their accuracy is not sufficient. Usually, US images have a lot of noise and artifacts and, especially in the case of breast LDR-PS treatment where various targets are difficult to discern solely based on image characteristics [3]. It is challenging to locate preoperatively identified targets quantitatively in the intraoperative physical situation because the organ/breast's shape changes between preoperative imaging and intraoperative situations [4], [5]. To improve the fidelity of breast brachytherapy and to improve target tracking accuracy during surgery, it is essential to use intraoperatively available data to accurately deform the preoperative model of the organ to simulate the deformation during surgery.

Finite Element Methods (FEM) are used for tissue simulation but complex FEMs are computationally expensive. Simulation Open Framework Architecture (SOFA) is an open-source framework that includes FEM models for medical applications, which reduces computational costs at the expense of accuracy [6]. Neural networks trained from synthetic FEM simulations can learn complex FEM behaviour, but there is a sim-to-real gap [7]. Fine-tuning with real-world data can improve the model, but requires a large amount of recorded data and cannot be done online [8], [9]. To resolve the problem of simulation-reality mismatch, researchers have used deep learning models to simulate deformable objects using visual information from deformed tissue during surgery [10]–[12]. However, this method has limitations as it cannot provide information on the internal structure of tissues/organs, which is necessary for many surgical procedures [12].

sim-to-real approaches have been used in physics-based simulators in order to mitigate model mismatches by incorporating real-world data. Several methods have been proposed to address the sim-to-real gap, primarily based on two main categories: 1) simulation parameter inference using real data to make simulations realistic [13]–[17], and 2) residual models that an auxiliary model attempts to rectify the sim-to-real mismatch [9], [18].

1) Parameter inference: The authors of [15] formulate state space equations of deformable objects using distribution representation, which allows for better incorporation of state observations in Bayesian parameter estimation. In [16] differentiable point cloud sampling and differentiable simulation are

used to perform simulation parameter inference. The parameter inference approaches estimates model parameters offline using recorded trajectories, which can be computationally challenging and limited in usefulness for real-time registration in surgical applications [13]–[17].

2) Residual models: Combining the base model either a physics-based or off-line learned model with a residual model is also beneficial to resolve the sim-to-real gap [18]–[21]. For complex deformable objects, online learning a residual model which is data-efficient has been a challenge in the literature. [18] proposed a linear residual model based on local Jacobian estimation to rectify the out-put error of a GNN when it is used to predict the state of a cable. In order to make a jacobian prediction, it requires access to all states of the system, and the deformation must be small. Because of these two drawbacks, local Jacobian learning is not applicable to the entire mesh update in LDR-PS. The author in [22] developed an online iterative residual framework to update the output of Position-Based Dynamics (PBD) simulation based on 3D visual perception. However, in [22], the proposed framework has not been tested for the ability to predict deformations of the internal points of the tissues. In [23], the KF-ADMM method registers the simulation output of a physics-based simulator with real data taken from the surface of the tissue using a Kalman filtering framework. In this work instead of learning an explicit residual model, the effect of unmodeled dynamics terms is formulated as zero-mean Gaussian distribution and overcome using a Kalman filtering framework. zero-mean Gaussian is a restrictive assumption due to the nonlinearity of tissue dynamics.

A. Objective and Contributions

This paper focuses on addressing the challenge of sim-to-real transfer of entire deformable objects using an online residual model. The proposed method refines the prediction obtained from a physics-based tissue simulator during the online phase, without requiring pre-collected real-world data to learn the residual model. Additionally, the method can compensate for large prediction errors. As a solution to this challenge, we propose a sim-to-real transfer framework that utilizes deep learning networks for representing deformed meshes in a lower-dimensional space, as well as an ensemble method for assimilation of data in order to bring real-world partial measurements of the simulator’s output together with the simulator’s predictions. The current paper contributions are compared with the various sim-to-real methods in Tabel I.

The proposed sim-to-real framework utilizes a variational auto-encoder (VAE) with graph-neural networks in the encoder and decoder parts to learn the probability distributions of low-dimensional latent variables. In this paper, VAE is preferred over GAN, since it is able to learn the underlying probability distribution of the data which is necessary in data assimilation task and it is more suitable for small datasets. Based on real measurements taken from the surface of the tissue, this framework uses a data-assimilation method to derive latent-space variables. The GNN-VAE is trained on a synthetic dataset of deformed meshes obtained from various FEM simulations. The

proposed method updates the prior distribution of latent variables iteratively based on the difference between the simulated mesh and the real tissue’s surface deformations observed by external trackers during the operation. An ensemble smoother with multiple data assimilation (ES-MDA) is used to update the prior distributions, and the decoder is used to reconstruct the updated deformed mesh after each iteration of ES-MDA.

In summary, the contributions of this study are as follows:

- 1) Real-time registration of deformed meshes derived from physics-based tissue simulators (such as FEM) is accomplished through deep learning and data assimilation methods.
- 2) With the data-assimilation method integrated into latent space, the entire mesh structure is updated efficiently in a timely manner.
- 3) An ensemble smoother with multiple data assimilation (ES-MDA) is used to implement the data assimilation and integrate discrete data points from the tissue surface into simulation results.
- 4) To enhance the time efficiency of ES-MDA, the standard ensemble generation and forecast steps are replaced with the forward step of GNN-VAE networks. This substitution enables faster real-time registration of deformed meshes obtained from physics-based tissue simulators.
- 5) The proposed sim-to-real framework does not require observation of the entire tissue since it only requires a few discrete measurements on its surface. The latent space does not require separate encoding of measurements.

The remaining of the paper is organized as follows: In Section II, our proposed sim-to-real framework and all its component will be explained comprehensively. In Section III, the training process of GNN-VAE will be explained and the accuracy of the method in the simulation will be investigated. Later, in Section IV, the performance of the method will be demonstrated experimentally and a conclusion in Section V will complete the paper.

II. THE PROPOSED SIM-TO-REAL FRAMEWORK

The flowchart of the proposed method is shown in Fig. 1. In the proposed method, the output of a physics-based deformable object simulator, which is chosen to be the Finite Element Method (FEM) in this paper, is the input to the sim-to-real module. FEM is implemented using the FEBio package, and it can be replaced with any other tissue simulator. The output of the sim-to-real module can be used as the input for FEM at the next time step.

The sim-to-real module can be seen as a Data Assimilation (DA) module that approximates the true states/parameters of the physical system by combining real-world observations with a theoretical model. Ensemble-based methods are among the most successful and efficient techniques currently available for DA. To alleviate the burden of high-dimensional calculations that would be necessary, e.g., when updating a large mesh model of soft tissue, DA must be performed in a lower-dimensional space that still encapsulates the principal features

TABLE I: Summary of characteristics of sim-to-real methods.

Paper	Sim-to-real approach	Offline-learning	Online-update	Linear/Non-Linear	Application
[15], [16]	Parameter inference	✓	×	Non-linear	General
[18]	Residual models	✓	✓	Linear update (Small deformation)	Cabel (full state observable)
[22]	Residual models	×	✓	Linear update	Tissue Surface
[23]	Residual models	×	✓	Linear update	Entire tissue
Proposed method	Residual models	✓	✓	Non-linear update	Entire tissue

of the original mesh. This lower-dimensional space is called the latent space.

In ensemble-based DA methods, hundreds of realizations of states/parameters must be generated and fed to the FEM to estimate the prior distributions of states/parameters at each time step. Instead of generating hundreds of ensembles to replicate the probabilistic characteristics of estimation at each step, we propose the use of variational auto-encoders (VAE). The use of VAE can significantly reduce the computational cost of generating ensembles.

In summary, the sim-to-real module consists of two steps. In the first step, the distributions of the latent variable associated with the output mesh of FEM simulation are computed using the GNN-VAE network. The GNN is used to encode the topology of the mesh as a graph, while the VAE is used to learn a low-dimensional probabilistic representation of the graph that captures the variation in the shape of the mesh. Once the GNN-VAE model is trained, it can be used to generate new meshes with different topologies from the input mesh. In the second step, the ES-MDA method incorporates the real-world measurements of the actual tissue surface deformations at each time step to update the prior distributions of latent variables and get the posterior distributions of the latent variables. Finally, the mean of posterior distribution as the most probable combination of the latent variable is selected and by feeding it to the pre-trained decoder part of GNN-VAE, an updated simulated mesh compatible with real-world measurements is constructed.

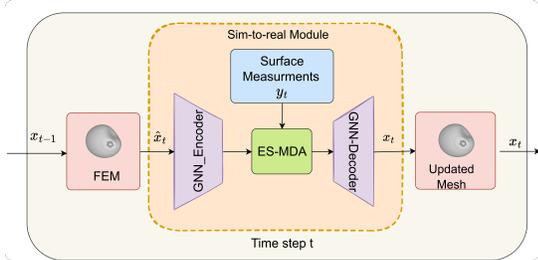


Fig. 1: The flowchart of the proposed sim-to-real framework.

In the next section, Graph-neural network Variational Autoencoders (GNN-VAE) are briefly explained and then the applied DA for sim-to-real using the Ensemble Smoother with Multiple Data Assimilation (ES-MDA) is explained.

A. Graph-based Variational Auto-Encoders

The VAE structure is described in this section. Next, the application of Graph Neural Networks is discussed for the

extraction of deformation features from deformed meshes. Finally, a GNN-VAE network is designed by combining these two structures.

1) *Variational Auto-Encoders*: Autoencoders are a class of unsupervised neural networks that are widely used for representation learning and dimension reduction. An autoencoder consists of two components: an encoder and a decoder. The encoder aims to extract low-dimensional latent features z from the high-dimensional input data x , whereas the decoder aims to recover the predicted input data \hat{x} from the latent features while minimizing the reconstruction error. In the VAE structure, a function $E_\theta(x)$ receives a sample from $\mathbf{x} \sim p(\mathbf{x})$ and generates a distribution of latent-variable \mathbf{z} , then a function $D_\theta(z)$ which receives a random argument $\mathbf{z} \sim p(\mathbf{z})$ and generates a sample from learned distribution $\hat{\mathbf{x}} \sim p_\theta(\mathbf{x} | \mathbf{z})$.

In the proposed framework, x represents a realization of the deformed mesh (i.e., the output of FEM simulation as it is shown in Fig. 1). Training samples, x_i , are available from generated dataset in terms of a patient-specific mesh whose mechanical parameters vary in a predefined range and undergo various force excitations. Thanks to having a generative model, it is easy to generate new deformed meshes that are distinguishable from the initial output of FEM simulations. After training, by sampling from the multi-variable learned distribution in the latent space, various meshes can be generated. This probabilistic distribution is the prior distribution of the FEM simulation. The objective is to update the prior distribution of latent variable using a data-assimilation method based on measurements coming from the real tissue and get the posterior distribution of latent variable and then sample the latent variables with the highest possibilities (see Fig. 1).

2) *Graph Neural Networks*: In the VAE structure, $E_\theta(x)$ and $D_\theta(z)$ are neural networks composed of layers compatible with data structures. While images and time series belong to Euclidean domains, tetrahedral meshes belong to irregular and non-Euclidean domains that can be represented with graphs. It is not possible to directly apply ordinary 2D or 3D convolution networks to mesh data due to the irregularities in local structures in meshes (varying vertex degrees, varying sampling densities, etc.). Graph neural networks are designed to extract information from graph data structures. An autoencoder based on spatially variable convolution kernels has been proposed in [24], where each vertex has its own convolution kernel. Based on a global kernel weight basis, a vertex-specific kernel is estimated. As the training process progresses, the global kernel weight basis, as well as a sampling function for each individual kernel, is learned. In irregular mesh connections, the spatially-

varying convolutions layer provides efficient means of capturing the spatially-varying contents. In this paper, the spatially-varying convolution layer and pooling layer introduced in [24] are used to build the encoder and decoder of VAE.

a) *Fully Convolutional Graph Layer*: In a convolutional layer, the input data is $x \in \mathcal{R}^{V \times d}$ where V is the number of vertices, and d is the dimension of input data, and produces output data $y \in \mathcal{R}^{V \times d'}$ where d' is the dimension of the output data. A schematic of the convolution operator is shown in Fig. 2a. The convolution operator for each vertex of a graph can be calculated using

$$y_i = \sum_{\mathbf{x}_{i,j} \in \mathcal{N}(i)} \mathbf{W}_j^T \mathbf{x}_{i,j} + \mathbf{b}. \quad (1)$$

Due to the uneven distribution of vertices on a mesh, and the different connectivity between vertices, the same weighting schemes cannot be applied for each vertex. Each vertex should be able to determine its convolution weight freely. In [24], a discrete convolution kernel is defined with weights on a standard grid which are called Weight Basis as shown in Fig. 2b. The vertices of a local region of the mesh scatter within the grid. In (2), the weights at real vertices can be sampled from a Weight Basis via different functions from vertex to vertex.

$$\mathbf{W}_{i,j} = \sum_{k=1}^M \alpha_{i,j,k} \mathbf{B}_k \quad (2)$$

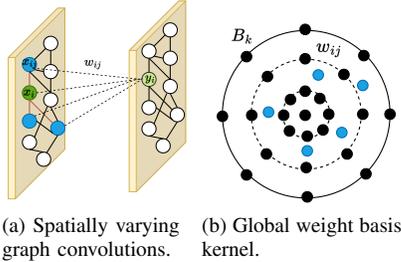


Fig. 2: Graph convolution and global weight basis kernel introduced in [24].

b) *Pooling Graph Layer*: In an arbitrary graph, the vertices can be distributed quite unevenly within the kernel radius, and using max or average pooling does not perform well. A pooling layer is introduced in [24], which applies Monte Carlo sampling for feature aggregation as shown in Fig. 3. In the pooling layer, the stride is 2 and radius is 1. The output feature of aggregated input nodes can be calculated using

$$y_i = \sum_{j \in \mathcal{N}(i)} \rho'_{i,j} \mathbf{x}_{i,j}, \quad \rho'_{i,j} = \frac{|\rho_{i,j}|}{\sum_{j=1}^{E_i} |\rho_{i,j}|} \quad (3)$$

where $\rho_{i,j} \in \mathbb{R}$.

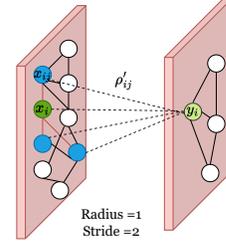


Fig. 3: Pooling layer on a graph data with radius=1 and stride=2.

B. The proposed sim-to-real module: Data-Assimilation with GNN-VAE

This paper aims to propose an approach for updating the output mesh of a finite element model (FEM) at each time step, based solely on data from the object surface, without replacing FEM with the complicated temporal deep network. While the Ensemble Kalman Filter (EnKF) is an effective method for sequential data assimilation of non-linear systems, it requires learning the temporal dependency using a complex network, making it unsuitable for our purpose. Unlike EnKF, Ensemble Smoother (ES) does not assimilate data sequentially in time. Instead, ES computes a global update by simultaneously assimilating all data available. However, the accuracy of one-step ES is limited due to the large step size of ensemble updates. To address this issue, the authors suggest employing Ensemble Smoother with Multiple Data Assimilation (ES-MDA) [25], which uses an iterative approach based on measurements at the current step to estimate the values of unknown parameters. ES-MDA involves assimilating the same data multiple times, with the covariance of measurement errors multiplied by the number of assimilations, resulting in improved accuracy.

In the proposed method, the unknown parameters are low-dimensional latent-space variables, and the relationship between these parameters and the observations is represented by a forward model, which is the pre-trained decoder network of GNN-VAEs. The ES-MDA method updates the estimates of the unknown parameters iteratively using the available observations and the forward model to refine the estimates at each step. This can be used to accurately estimate the values of the latent-space variables that are compatible with the observations.

This sim-to-real module integrates ES-MDA data assimilation with GNN-VAE. The following is a summary of the detailed steps of the proposed sim-to-real module:

- 1) At each time step, the output of FEM simulation is fed into the sim-to-real module as it is shown in Fig. 1 and Fig. 4.
- 2) Initialization step of the sim-to-real module: In traditional ES-MDA, initial ensembles of parameters must be defined based on prior data. However, in the proposed method, prior ensembles are generated by sampling from a normal distribution in the latent space, $\mathbf{z}_i \sim p(\mathbf{z})$, where $p(\mathbf{z}) = \mathbb{E}(\mathbf{x}_i)$, and \mathbb{E} denotes the encoder network of GNN-VAE. The number of iterations must then be determined, and the next two steps are repeated

for that number of iterations.

- 3) Forecast step of the sim-to-real module: The ensemble realization i is used as input into the forward model, the decoder network of trained GNN-VAE, which produces an output mesh. The surface points from the output mesh are selected to produce an ensemble of model prediction \mathbf{y}_i at each measurement location, $\mathbf{y}_i^n = \mathbb{D}(\mathbf{z}_i^n)$, where \mathbb{D} denotes the forward model which is the decoder network of GNN-VAE, i is the realization index and n is the iteration index. The forecast step is shown in Fig. 4.
- 4) Update step of the sim-to-real module: Latent-space realizations are updated at each time step using a single set of measurement from that time step. To enable iterative data assimilation based on one measurement, the measurement vector is disturbed at each iteration using a noise vector multiplied by an inflated covariance error matrix. Inflating the measurement error covariance matrix dampens extreme changes in the model during early iterations. The difference between the disturbed measurement vector and ensemble predictions is then calculated and weighted based on the covariance matrices to maximize the likelihood of ensemble prediction. The update rule can be expressed mathematically using the following equation [25]

$$\mathbf{z}_i^{n+1} = \mathbf{z}_i^n + \mathbf{C}_{\mathbf{z}\mathbf{y}}^n (\mathbf{C}_{\mathbf{y}\mathbf{y}}^n + \alpha_n \mathbf{C}_{\mathbf{d}})^{-1} (\mathbf{d}_{\text{obs}} + \sqrt{\alpha_n} \mathbf{C}_{\mathbf{d}}^{1/2} \epsilon_i^n - \mathbf{y}_i^n) \quad (4)$$

where $\mathbf{C}_{\mathbf{d}}$ is the user-defined covariance matrix and ϵ_i^n is the observation error at iteration n , which is drawn from a Gaussian distribution $\mathcal{N}(0, I_{N_d})$ which N_d is the number of observations. α_n is a coefficient that, at each iteration n , inflates the measurement error and its covariance matrix. Values are selected in a decreasing order; in this way, the magnitude of the updates for the first iterations, when there might be a large misfit between predictions and observations, will be smaller to reduce the magnitude of initial updates; also, the coefficients α_n must satisfy $\sum_{n=1}^{N_a} \frac{1}{\alpha_n} = 1$ conditions, where N_a is the total number of iterations.

$\mathbf{C}_{\mathbf{z}\mathbf{y}}^n$ is the cross-covariance matrix between latent-space variables and surface point predictions and $\mathbf{C}_{\mathbf{y}\mathbf{y}}^n$ is the autocovariance matrix of surface point predictions. They are computed from the ensemble at each iteration n using

$$\begin{aligned} \mathbf{C}_{\mathbf{z}\mathbf{y}}^n &= \frac{1}{N_e - 1} \sum_{i=1}^{N_e} (\mathbf{z}_i^n - \bar{\mathbf{z}}) (\mathbf{y}_i^n - \bar{\mathbf{y}})^T \\ \mathbf{C}_{\mathbf{y}\mathbf{y}}^n &= \frac{1}{N_e - 1} \sum_{i=1}^{N_e} (\mathbf{y}_i^n - \bar{\mathbf{y}}) (\mathbf{y}_i^n - \bar{\mathbf{y}})^T \end{aligned} \quad (5)$$

where N_e is the total number of ensemble realizations, $\bar{\mathbf{z}}$ is the ensemble mean of the latent-space variables and $\bar{\mathbf{y}}$ is the ensemble mean of the surface point predictions.

All steps are elaborated in Algorithm 1.

Algorithm 1: The proposed sim-to-real algorithm.

Input: $N_a, N_e, \mathbf{C}_{\mathbf{d}}, \alpha$, Pre-trained GNN-VAE, and Input mesh from FEM: M_t

Output: Updated mesh at time t: M_t^*

```

1 for time step t do
2    $M_t \leftarrow$  Output of FEM simulation at time t.
   Calculate the prior distribution at time t:
    $p_t(z) \leftarrow \mathbb{E}(M_t)$ 
3   Sample  $N_e$  ensembles from  $p_t(z)$ .
4   while  $N \leq N_a$  do
5     Forecast step:
6      $\mathbf{y}_i^n = \mathbb{D}(\mathbf{z}_i^n)$ 
7     Update step:
8      $\mathbf{z}_i^{n+1} = \mathbf{z}_i^n + \mathbf{C}_{\mathbf{z}\mathbf{y}}^n (\mathbf{C}_{\mathbf{y}\mathbf{y}}^n + \alpha_n \mathbf{C}_{\mathbf{d}})^{-1}$ 
        $(\mathbf{d}_{\text{obs}} + \sqrt{\alpha_n} \mathbf{C}_{\mathbf{d}}^{1/2} \epsilon_i^n - \mathbf{y}_i^n)$ 
9     where covariance can be calculated using 5.
10    end
11    Create the updated mesh:
12     $M_t^* = \mathbb{D}(\text{mean}(z))$ 
13  end
```

III. SIMULATION RESULTS

The target application of this paper is breast surgeries and the designed GNN-VAE for the breast mesh $x \in \mathcal{R}^{4223 \times 3}$ is depicted in Fig. 5. As it is shown in Fig. 5, the encoder consists of 3 convolution layers and 3 pooling layers. Each convolution layer has stride = 2 and radius = 1. The details regarding the GNN-VAE hyperparameters are discussed in Table II.

TABLE II: Details of the GNN-VAE layers.

	Layer	Layer output
Encoder	Convolution ($s = 2, r = 1, f = 32$)	$x \in \mathcal{R}^{4223 \times 32}$
	Pooling ($s = 2, r = 1$)	$x \in \mathcal{R}^{811 \times 32}$
	Convolution ($s = 2, r = 1, f = 64$)	$x \in \mathcal{R}^{811 \times 64}$
	Pooling ($s = 2, r = 1$)	$x \in \mathcal{R}^{82 \times 64}$
	Convolution ($s = 2, r = 1, f = 128$)	$x \in \mathcal{R}^{82 \times 128}$
	Pooling ($s = 2, r = 1$)	$x \in \mathcal{R}^{6 \times 128}$
	Reshape	$x \in \mathcal{R}^{1 \times 768}$
Decoder	Fully connected (26)	$x \in \mathcal{R}^{1 \times 26}$
	Fully connected (768)	$x \in \mathcal{R}^{1 \times 768}$
	Reshape	$x \in \mathcal{R}^{6 \times 128}$
	Un-Pooling ($s = 2, r = 1$)	$x \in \mathcal{R}^{82 \times 128}$
	Trans-Convolution ($s = 2, r = 1, f = 64$)	$x \in \mathcal{R}^{82 \times 64}$
	Un-Pooling ($s = 2, r = 1$)	$x \in \mathcal{R}^{811 \times 64}$
	Trans-Convolution ($s = 2, r = 1, f = 32$)	$x \in \mathcal{R}^{811 \times 32}$
Un-Pooling ($s = 2, r = 1$)	$x \in \mathcal{R}^{4223 \times 32}$	
Trans-Convolution ($s = 2, r = 1, f = 3$)	$x \in \mathcal{R}^{4223 \times 3}$	

To train the GNN-VAE on simulated data, a dataset consisting of 10,000 deformed meshes obtained from FEM simulations is used. To generate the dataset, the module of elasticity of tissue is varied in the range of 10-60 Kpa, while the external force application points and the force magnitudes are changed randomly. The training set consists of 80%, the test set contains 10%, and the validation set contains 10% of samples.

The accuracy of the trained network on the test data is 0.024 cm and on the training data is 0.0205 cm. We selected ten random meshes from the test dataset as initial meshes in

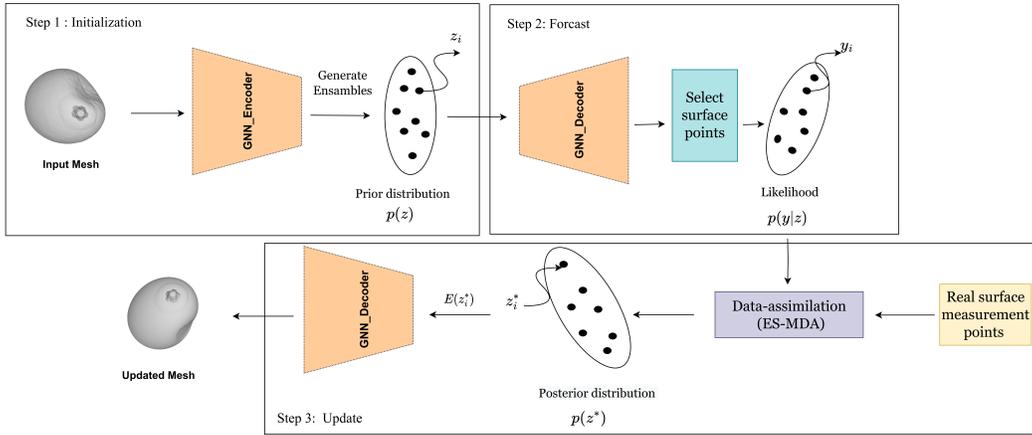


Fig. 4: The flowchart of sim-to-real module.

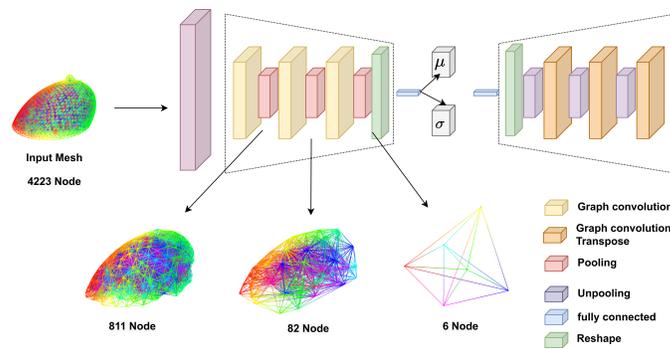


Fig. 5: GNN-based VAE structure.

the sim-to-real framework, as well as ten different random meshes from the test data set as target meshes associated with each of these initial meshes to determine the effect of the number of ensembles, N_e , number of update steps N_a , and measurement covariance matrix C_d on the accuracy of sim-to-real registrations.

In Fig. 6, $N_a = 5$ is fixed and the effect of N_e and C_d are investigated. Decreasing C_d from 0.1 to 0.001 decreases the average MSE error of the ten meshes. Also, it is shown in Fig. 6 that there is not a substantial difference between $C_d = 0.01$ and $C_d = 0.001$. Furthermore, choosing $N_e = 10$ results in inferior performance in comparison to $N_e = 50$ and $N_e = 100$, but once again the difference between $N_e = 50$ and $N_e = 100$ is not tangible.

In Fig. 7, $C_d = 0.001$ is fixed and the effect of N_e and N_a are investigated. According to Fig. 7, increasing N_a does not affect the final accuracy of the sim-to-real registration. However, choosing N_a to be less than 5 deteriorates the performance of the framework.

Fig. 8 shows a series of mesh updates between an initial mesh and a ground truth mesh with $N_a = 5$, $N_e = 100$, and $C_d = 0.001$. As it is shown in Fig. 8, there is an initial error of 0.73 cm between the initial mesh and the ground truth (GT) mesh, which decreases to 0.054 cm at the end of sim-to-real iterations.

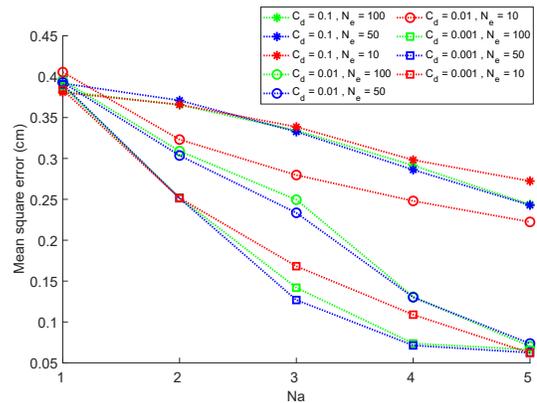


Fig. 6: An investigation of the effect of numeric parameters C_d , and N_e on sim-to-real framework accuracy.

IV. EXPERIMENTAL STUDY AND RESULTS

In this section, the performance of the proposed sim-to-real module is validated in tissue deformation prediction. Experiments on phantom tissue have been conducted in which markers located on the tissue surface are utilized to track the tissue surface movements. The experimental setup shown in Fig. 9 was built. An Aurora electromagnetic (EM) tracker with a Planar 20-20 V2 Field Generator (NDI Europe GmbH, Radolfzell, Germany) was utilized to track the 3D position of a magnetic sensor which was buried inside the tissue

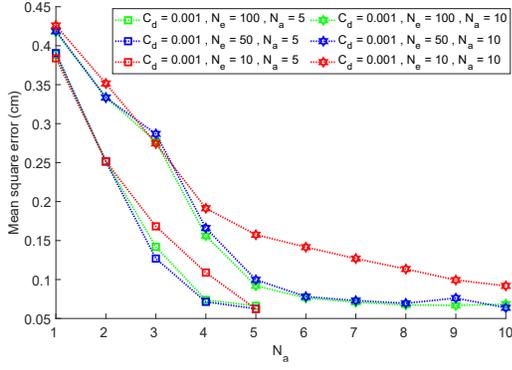


Fig. 7: An investigation of the effect of numeric parameters N_e , and N_a on sim-to-real framework accuracy.

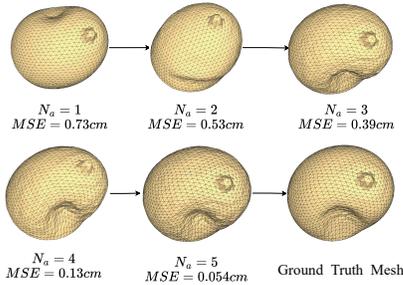


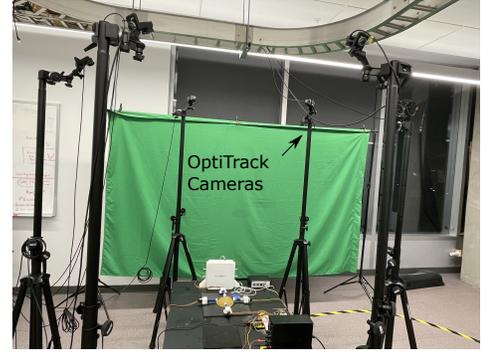
Fig. 8: sim-to-real updates between an initial mesh and a ground truth mesh. The MSE error is calculated between each mesh at each step and the GT mesh.

phantom as shown in Fig. 9. An Optic-track motion capture system with six cameras was used to track $4mm$ optic facial markers. The surface of the phantom was marked with 15 optic facial markers. The motion capture system can track facial markers with sub-millimetre accuracy after calibration. The force excitation are linear actuators pushing the breast phantom as shown in Fig. 9.

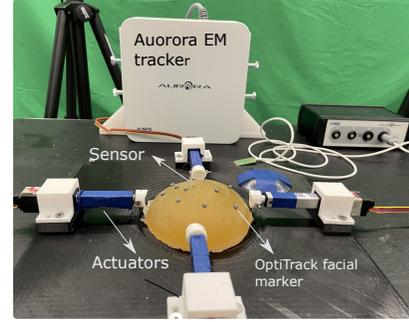
In the experiment, we will investigate how much the proposed sim-to-real module can update FEM simulation using a few measurements from the tissue surface and the performance is compared with another registration method called KF-ADMM from [23].

FEM with a Neo-Hookean material model with $E = 10Kpa$ is used for modelling the tissue deformation as it is being manipulated by linear actuators in Fig. 9. ADMM tissue simulation method uses the same mechanical parameters as FEM, i.e., the Neo-Hookean material model with $E = 10Kpa$ in the ADMM solver. The tissue has been deformed in two scenarios. In scenario 1, actuator number one pushes the phantom along the x-axis, and in scenario 2, actuators simultaneously push the phantom along the x-axis and y-axis.

The absolute error between the EM sensor measurements and those of the FEM simulation, as well as the revised predicted trajectory generated by the proposed sim-to-real method and KF-ADMM method, can be seen in Fig. 10 for scenario 1 and Fig. 11 for scenario 2. Based on Fig. 10, the postponed sim-to-real module can reduce prediction error more than KF-ADMM especially when the deformation is extreme, the proposed method shows better performance in



(a) The cameras' configuration.



(b) Setup details including EM sensor, optic markers, phantom and actuators.

Fig. 9: Experiment setup.

reducing the prediction error. In Fig. 11, the deformations are along two directions, and tissue deformation is more extreme than in scenario 1, the proposed method is more effective at reducing prediction error than KF-ADMM. In the current set of experiments, sim-to-real module reduced the FEM prediction error by 72% on average and performed 45% better than KF-ADMM.

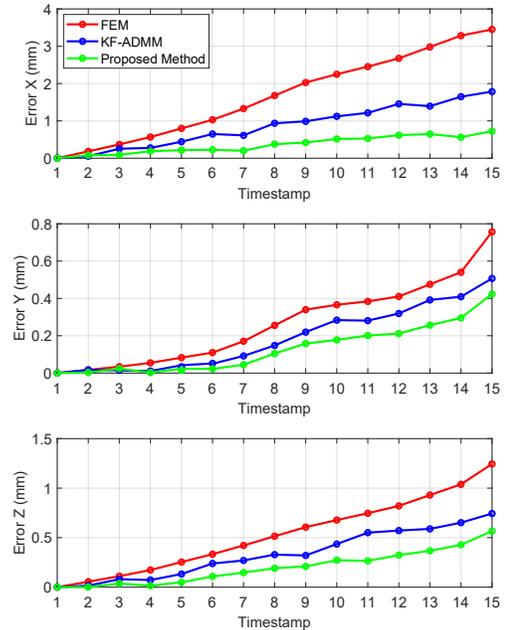


Fig. 10: Prediction error before and after registration based on the proposed method and KF-ADMM in scenario 1.

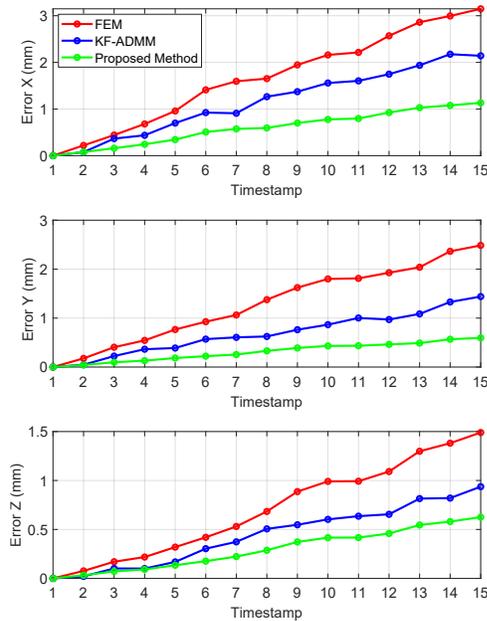


Fig. 11: Prediction error before and after registration based on the proposed method and KF-ADMM in scenario 2.

V. CONCLUSION

This article addresses the challenge of registering high-dimensional tissue deformation models from simulation to reality. In this study, a novel sim-to-real module was developed for registering a physics-based tissue simulation's output (here FEM was applied) to real measurements of deformed tissue. The proposed sim-to-real modules include graph-based variational auto-encoders (GNN-VAE) and an ensemble smoother with multiple data assimilation (ES-MDA). To solve the problem of updating a high-dimensional mesh in real-time for tissue deformation modelling, it integrates the generative auto-encoder networks for learning simulation-data distributions and the data-assimilation methods like ES-MDA for updating the learned distributions with real measurements. The GNN-VAE is trained on FEM simulation data and does not require retraining. Sim-to-real reduces the registration error more efficiently than KF-ADMM in extreme deformations, according to experiments. The method will be tested for more complex tissue manipulation tasks in the future.

REFERENCES

- [1] E. E. Deurloo, K. G. Gilhuijs, L. J. S. Kool, and S. H. Muller, "Displacement of breast tissue and needle deviations during stereotactic procedures," *Investigative radiology*, vol. 36, no. 6, pp. 347–353, 2001.
- [2] J. op den Buijs, M. Abayazid, C. L. de Korte, and S. Misra, "Target motion predictions for pre-operative planning during needle-based interventions," in *2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. IEEE, 2011, pp. 5380–5385.
- [3] J. Michael, D. Morton, D. Batchelar, M. Hilt, J. Crook, and A. Fenster, "Development of a 3d ultrasound guidance system for permanent breast seed implantation," *Medical Physics*, vol. 45, no. 8, pp. 3481–3495, 2018.
- [4] M. Afshar, J. Carriere, T. Meyer, R. Sloboda, S. Husain, N. Usmani, W. Liu, and M. Tavakoli, "Autonomous ultrasound scanning to localize needle tip in breast brachytherapy," in *2020 International Symposium on Medical Robotics (ISMR)*. IEEE, 2020, pp. 202–208.

- [5] M. Afshar, J. Carriere, T. Meyer, R. S. Sloboda, S. Husain, N. Usmani, and M. Tavakoli, "A model-based multi-point tissue manipulation for enhancing breast brachytherapy," *IEEE Transactions on Medical Robotics and Bionics*, vol. 4, no. 4, pp. 1046–1056, 2022.
- [6] F. Faure, C. Duriez, H. Delingette, J. Allard, B. Gilles, S. Marchesseau, H. Talbot, H. Courtecuisse, G. Bousquet, I. Peterliik, and S. Cotin, "SOFA: A Multi-Model Framework for Interactive Physical Simulation," in *Soft Tissue Biomechanical Modeling for Computer Assisted Surgery*, ser. Studies in Mechanobiology, Tissue Engineering and Biomaterials, Y. Payan, Ed. Springer, Jun. 2012, vol. 11, pp. 283–321. [Online]. Available: <https://hal.inria.fr/hal-00681539>
- [7] A. Mendizabal, E. Tagliabue, J.-N. Brunet, D. Dall'Alba, P. Fiorini, and S. Cotin, "Physics-based deep neural network for real-time lesion tracking in ultrasound-guided breast biopsy," in *Computational Biomechanics for Medicine*. Springer, 2019, pp. 33–45.
- [8] A. Kloss, S. Schaal, and J. Bohg, "Combining learned and analytical models for predicting action effects," *arXiv preprint arXiv:1710.04102*, vol. 11, 2017.
- [9] A. Allevato, E. S. Short, M. Pryor, and A. Thomaz, "Tunenet: One-shot residual tuning for system identification and sim-to-real robot task transfer," in *Conference on Robot Learning*. PMLR, 2020, pp. 445–455.
- [10] J. Lu, A. Jayakumari, F. Richter, Y. Li, and M. C. Yip, "Super deep: A surgical perception framework for robotic tissue manipulation using deep learning for feature extraction," in *2021 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2021, pp. 4783–4789.
- [11] J. Song, J. Wang, L. Zhao, S. Huang, and G. Dissanayake, "Dynamic reconstruction of deformable soft-tissue with stereo scope in minimal invasive surgery," *IEEE Robotics and Automation Letters*, vol. 3, no. 1, pp. 155–162, 2017.
- [12] Y. Yang, J. A. Stork, and T. Stoyanov, "Learning to propagate interaction effects for modeling deformable linear objects dynamics," in *2021 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2021, pp. 1950–1957.
- [13] F. Ramos, R. C. Possas, and D. Fox, "Bayessim: adaptive domain randomization via probabilistic inference for robotics simulators," *arXiv preprint arXiv:1906.01728*, 2019.
- [14] J. Liang, S. Saxena, and O. Kroemer, "Learning active task-oriented exploration policies for bridging the sim-to-real gap," *arXiv preprint arXiv:2006.01952*, 2020.
- [15] R. Antonova, J. Yang, P. Sundaresan, D. Fox, F. Ramos, and J. Bohg, "A bayesian treatment of real-to-sim for deformable object manipulation," *IEEE Robotics and Automation Letters*, vol. 7, no. 3, pp. 5819–5826, 2022.
- [16] P. Sundaresan, R. Antonova, and J. Bohg, "Diffcloud: Real-to-sim from point clouds with differentiable simulation and rendering of deformable objects," *arXiv preprint arXiv:2204.03139*, 2022.
- [17] A. Longhini, M. Moletta, A. Reichlin, M. C. Welle, D. Held, Z. Erickson, and D. Kragic, "Edo-net: Learning elastic properties of deformable objects from graph dynamics," *arXiv preprint arXiv:2209.08996*, 2022.
- [18] C. Wang, Y. Zhang, X. Zhang, Z. Wu, X. Zhu, S. Jin, T. Tang, and M. Tomizuka, "Offline-online learning of deformation model for cable manipulation with graph neural networks," *IEEE Robotics and Automation Letters*, vol. 7, no. 2, pp. 5544–5551, 2022.
- [19] A. Zeng, S. Song, J. Lee, A. Rodriguez, and T. Funkhouser, "Tossing-bot: Learning to throw arbitrary objects with residual physics," *IEEE Transactions on Robotics*, vol. 36, no. 4, pp. 1307–1319, 2020.
- [20] M. Alakujala, G. Dulac-Arnold, J. Mairal, J. Ponce, and C. Schmid, "Residual reinforcement learning from demonstrations," *arXiv preprint arXiv:2106.08050*, 2021.
- [21] C. Chi, B. Burchfiel, E. Cousineau, S. Feng, and S. Song, "Iterative residual policy: for goal-conditioned dynamic manipulation of deformable objects," *arXiv preprint arXiv:2203.00663*, 2022.
- [22] F. Liu, Z. Li, Y. Han, J. Lu, F. Richter, and M. C. Yip, "Real-to-sim registration of deformable soft tissue with position-based dynamics for surgical robot autonomy," in *2021 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2021, pp. 12 328–12 334.
- [23] M. Afshar, J. Carriere, H. Rouhani, T. Meyer, R. S. Sloboda, S. Husain, N. Usmani, and M. Tavakoli, "Accurate tissue deformation modeling using a kalman filter and admm-based projective dynamics," *IEEE/ASME Transactions on Mechatronics*, 2022.
- [24] Y. Zhou, C. Wu, Z. Li, C. Cao, Y. Ye, J. Saragih, H. Li, and Y. Sheikh, "Fully convolutional mesh autoencoder using efficient spatially varying kernels," *Advances in Neural Information Processing Systems*, vol. 33, pp. 9251–9262, 2020.
- [25] A. A. Emerick and A. C. Reynolds, "Ensemble smoother with multiple data assimilation," *Computers & Geosciences*, vol. 55, pp. 3–15, 2013.



Mehrnoosh Afshar received the BSc degree in mechanical engineering from Amirkabir University of Technology (Tehran Polytechnic) in 2014 and the MSc degree in mechanical engineering from the Sharif University of Technology in 2016. She is currently pursuing a Ph.D. in Electrical Engineering at the University of Alberta. Her research interests involve modeling and application of theory of control and AI in medical application. Currently, she focuses on developing systems for Robot-assisted breast brachytherapy.



Tyler Meyer Meyer is now a medical physicist at the Tom Baker Cancer Centre in Calgary. He is also an adjunct associate professor at the U of C in the department of oncology, with a secondary appointment in the department of physics and astronomy. The primary focus of his research and clinical practice is brachytherapy, a classification of radiotherapy where radioactive seeds or sources are placed in or near the target area. It's commonly used in cancer treatment for prostate, cervical and other gynecological cancers. But Meyer is also doing

groundbreaking research on breast brachytherapy.



Ron S. Sloboda received the B.Sc. degree in physics from the University of Manitoba, Winnipeg, MB, Canada, in 1974, and the Ph.D. degree in physics, nuclear theory, from the University of Alberta, Edmonton, AB, Canada, in 1979. He is currently a Professor in the Department of Oncology, University of Alberta. His research interests include dosimetry and treatment planning for brachytherapy, including the design of clinical studies to obtain patient data and model based dose calculation.



Siraj Hussain Husain completed medical school at the Dalhousie University (Halifax, Nova Scotia) followed by residency training at the University of Alberta (Edmonton, Alberta). He completed a fellowship in Radiation Oncology at the University of Western Ontario (London, Ontario). His current clinical areas of focus include genitourinary and breast cancer. He has clinical and research interests in Brachytherapy and leads the implementation and training of HDR to new physicians at the Tom Baker Cancer Centre (TBCC). Dr. Husain is the

Clinical and Research Program Leader for Prostate Brachytherapy, Radiation Oncology, at the TBCC. Active areas of research include enhancement and progression of combined modality therapy (Prostate Brachy + External Beam Radiation Therapy), champion for intermediate and high risk prostate cancer “drafted protocol” for introduction of High Dose Rate (HDR) and research, continue to advocate for Breast Brachytherapy – with new research project to enhance and improve the procedure.



Nawaid Usmani main focus on research is in prostate brachytherapy. Dr. Usmani's main objective for this research is to characterize current brachytherapy techniques and identify strategies for improving this treatment. In addition to this research in prostate brachytherapy, Dr. Usmani is involved in a number of other research endeavours. His other research includes: Investigating the potential benefits of metformin in preventing metabolic complications of hormonal therapy and improving prostate cancer outcomes; Identifying new prognostic or predictive biomarkers in prostate cancer; Investigating the utility of magnetic resonant imaging and PET imaging in the management of prostate cancer; and Investigating the potential benefits of exercise in rectal cancer patients.



Mahdi Tavakoli is a Professor in the Department of Electrical and Computer Engineering, University of Alberta, Canada. He received his BSc and MSc degrees in Electrical Engineering from Ferdowsi University and K.N. Toosi University, Iran, in 1996 and 1999, respectively. He received his PhD degree in Electrical and Computer Engineering from the University of Western Ontario, Canada, in 2005. In 2006, he was a post-doctoral researcher at Canadian Surgical Technologies and Advanced Robotics (CSTAR), Canada. In 2007-2008, he was

an NSERC Post-Doctoral Fellow at Harvard University, USA. Dr. Tavakoli's research interests broadly involve the areas of robotics and systems control. Specifically, his research focuses on haptics and teleoperation control, medical robotics, and image-guided surgery. Dr. Tavakoli is the lead author of Haptics for Teleoperated Surgical Robotic Systems (World Scientific, 2008). He is a Senior Member of IEEE and an Associate Editor for IEEE/ASME Transactions on Mechatronics, Journal of Medical Robotics Research, IET Control Theory Applications, and Mechatronics.