2024 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)
October 14-18, 2024. Abu Dhabi, UAE

A Novel Approach for Precise Tissue Tracking in Breast Lumpectomy

Yeganeh Aliyari[1], Mehrnoosh Afshar[1], Ericka Wiebe[2],
Lashan Peiris [3], and Mahdi Tavakoli[1,4], *Senior Member, IEEE*

***Abstract*— One of the most common cancers among women is breast cancer which can be treated surgically in the early stages with a lumpectomy technique. In the context of breast lumpectomy procedures, accurately tracking tumours presents a critical challenge worsened by various sources of anatomical deformations, including breathing, tissue cutting, and ultrasound probe pressure. To address this, we explore how a realistic tissue deformation simulator can enhance the precision of locating internal targets by accurately assessing the deformation applied to a preoperative model of the breast, considering the distinct mechanical properties of both the breast tissue and the tumour within it. Our method uses advanced artificial intelligence techniques by combining a generative variation autoencoder (GNN-VAE) and an updating method called ensemble smoother with multiple data assimilation (ES-MDA), creating a dynamic model based exclusively on surface node data to update all nodes within the tissue. By leveraging a realistic tissue deformation simulator, our approach uses breast surface tracking to infer full tissue deformations. This makes the method compatible with various simulation tools and suitable for tissues with complex properties. The results indicate that the trained network has an accuracy of 0.014 cm with training data, and 0.026 cm with the testing portion of data, demonstrating precision in tumour localization and significantly improving upon current methods. This innovation can potentially enhance patient outcomes by making breast cancer surgery safer, less invasive, and more efficient.**

I. Introduction

Breast cancer continues to be one of the most frequently detected forms of cancer and the leading cause of cancer-related fatalities among women. Annually, it results in the loss of over 600,000 lives globally [1]. Screenings and awareness of risk factors are vital for early detection, which can be treated with surgery, radiation, and chemotherapy. Breast cancers that are in their early stages of development can be treated surgically by undergoing either a mastectomy (complete breast removal) or a lumpectomy (breast-conserving surgery removing only the tumour and adjacent tissue). To minimize morbidity including pain, infection, and impact on quality of life, lumpectomy is often preferred. Clear delineation of the tumour location is necessary for successful lumpectomy surgery. However, the breast is remarkably flexible and will deform based on the patient's positioning, tissue cutting, or the ultrasound probe pressure during scanning [2]. This complicates the locating and tracking of impalpable tumours during surgery. The ability to accurately identify and remove cancerous tissue while sparing healthy tissue is central to treatment efficacy and the cosmetic outcomes patients value [3].

A common method for locating tumours is wire localization. An ultrasound or mammogram scan is used to find the tumour location and place a wire percutaneously in the breast, with the distal wire segment adjacent to the tumour and the proximal wire segment outside the breast. To prevent any disturbance to the wire's positioning before surgery, patients must be cautious due to the wire's external component and the breast's deformable nature [4]. Moreover, there have been advancements in non-wire localization methods, which involve using small markers such as radioactive or magnetic seeds or radiofrequency identification tags. These markers are inserted into the tissue with the help of imaging techniques and then identified during surgery using corresponding detectors [5]. However, these methods are invasive and lead to patient discomfort [6].

Overall, despite advancements in wire and non-wire localization techniques, these methods often fall short of accurately navigating the complexities of tissue deformation during lumpectomy procedures. These limitations not only hinder the surgeon's ability to localize precisely and excise tumours but also increase the risk of incomplete tumour removal or excessive removal of healthy tissue, adversely affecting surgical outcomes [6]. This gap underscores the need for a novel approach to adapt to and account for tissue deformations in real-time, ensuring accurate tumour localization and improved surgical precision.

In light of this, the development of non-rigid registration algorithms based on a simulation of the biomechanical behaviour of breast tissue has become an active area of research in recent years [3]. This approach aims to accurately model and track the complex biomechanical behaviour of breast tissue during procedures. However, there are certain limitations associated with this approach, including establishing a realistic constitutive model for breast tissue, obtaining the elastic constants specific to that patient, establishing appropriate boundary conditions, or reducing the computational cost to levels considered clinically acceptable [7].

Finite element methods (FEMs) are commonly employed in biomechanics to model the deformation of soft tissues;

*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 Innovates

[1]Y. Aliyari, M. Afshar, and M. Tavakoli are with the Department of Electrical and Computer Engineering, University of Alberta, Edmonton, AB, Canada. aliyaris, afsharbo, mahdi.tavakoli@ualberta.ca

[2]E. Wiebe, Oncology, with the Department of Medicine and Dentistry University of Alberta, Edmonton, AB, Canada emwiebe@ualberta.ca

[3] L. Peiris, Surgery, with the Department of Medicine and Dentistry University of Alberta, Edmonton, AB, Canada peiris@ualberta.ca

[4]M. Tavakoli is with the Department of Biomedical Engineering, University of Alberta, Edmonton, AB, Canada

however, complex FEMs are computationally expensive [8]. Employing neural networks trained on synthetic FEM simulations can learn complex FEM behaviour. Even with complex models, a gap exists between simulation and real-world application [9]. Fine-tuning models with real-world data enhances their performance but necessitates substantial recorded data and offline processing [10]. To address the simulation-reality disparity, deep learning techniques have been utilized to simulate deformable objects based on visual cues from surgical procedures [11]. However, this approach lacks insight into tissue internal structures crucial for surgeries [12].

Various sim-to-real methods have been proposed, primarily focusing on simulation parameter inference from real data and residual models. A review of recent literature is summarized here for described approaches with varying benefits and limitations. Parameter inference involves estimating model parameters offline using recorded trajectories, posing challenges for real-time surgical applications [13]. Residual models are a crucial method for narrowing the gap between simulation and reality, achieved by integrating base models with auxiliary ones. However, the online learning process for complex deformable objects using residual models poses significant challenges. Previous approaches, such as linear residual models and iterative frameworks relying on 3D visual perception, have demonstrated these difficulties [14]. A recent paper [15] proposes a linear residual model using local Jacobian estimation to correct output errors in a Graph Neural Network (GNN) predicting cable states. This method requires access to all system states and small deformations, which limits its applicability. Liu *et al.* introduce an online iterative residual framework for updating Position-Based Dynamics (PBD) simulation outputs based on 3D visual perception, yet it has not been tested for predicting tissue internal point deformations [16]. Lastly, in [17] presents the KF-ADMM method, which aligns physics-based simulation outputs with real tissue data using Kalman filtering without explicitly learning a residual model. However, the assumption of a zero-mean Gaussian distribution may be restrictive due to tissue dynamics' nonlinearity.

This study aims to overcome the difficulty associated with tracking deformable tumours within a deformable breast with varying mechanical properties during lumpectomy by employing an online residual model. Predictions derived from a physics-based tissue simulator are enhanced via the proposed method during the online phase, eliminating the need for pre-gathered real-world data during the training phase. Moreover, the approach is capable of correcting substantial prediction inaccuracies. To address tracking a deformable target within a deformable tissue, and to enhance the simulation's resemblance to the real world, the framework initially undergoes training using less sophisticated datasets for tissue deformation and displacement. It also uses a more sophisticated simulated tissue model that is considered to represent real-world conditions as the ground truth. The simulations are conducted in advance, and the outcomes are then fed into a GNN-VAE to produce a probabilistic low-dimensional representation of the simulator's outputs. The derived low-dimensional distribution, referred to as the prior distribution, undergoes iterative updates by applying an ES-MDA approach, and the decoder is employed to reconstruct the updated deformed mesh.

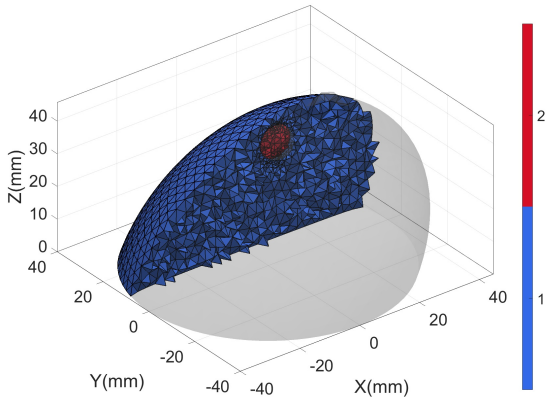

Fig. 1: Merged model showing both the breast and the tumour.

The remainder of the paper is organized as follows: Section II outlines the Developing a Comprehensive Dataset for Modeling Breast Tissue Deformation. Section III discusses the proposed GNN-VAE and ES-MDA frameworks. Simulations and results of the proposed method are provided in Section IV. Section V concludes the paper with remarks on potential future work.

II. Developing a Comprehensive Dataset for Modeling Breast Tissue Deformation

This section details the creation of a specialized dataset aimed at accurately modelling breast tissue deformations for improved tissue tracking. It is designed to enhance the simulation of tissue mechanics and form the foundation for our novel approach. The breast, along with the tumour inside it, is a heterogeneous structure and the finite element (FE) volume mesh was generated by applying TetGen [18] to the surface mesh files. Fig. 1 depicts the FE volume mesh, comprising 10859 nodes. Among these, the tumour is represented by 453 nodes, characterized by an elasticity modulus of $E = 20kPa$ and a Poisson's ratio of 0.2. In addition, the remaining nodes represent the healthy tissue, featuring an elasticity modulus of $E = 10kPa$ and a Poisson's ratio of 0.45 [19]. Tumours are significantly stiffer than normal breast tissues. This stiffening is attributed to collagen remodelling in the tumour tissues [19].

The boundary conditions were established, as depicted in Fig. 2. The base of the breast was set to be fixed, while the remaining nodes were left free to move. The red nodes indicate the locations where external forces can be applied, and we randomly vary these locations and the magnitudes of forces. For each deformation, a force direction is established as the surface normal at the chosen point, with a random angle introduced to accommodate variations in tool compression angles. This process is repeated until the desired

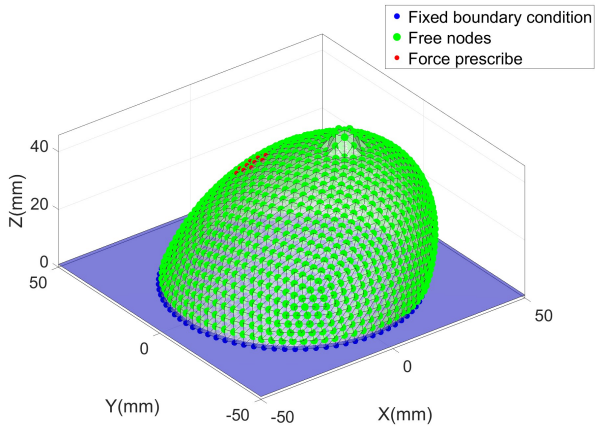

Fig. 2: Boundary conditions, free nodes, and the nodes where loads were applied.

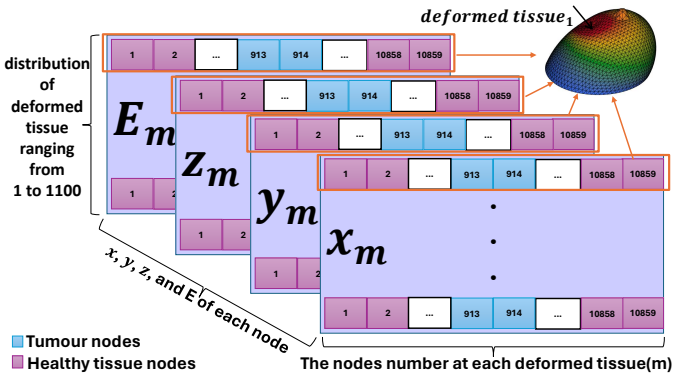

Fig. 3: The dataset structure.

number of samples is obtained. By allowing for variations in the angle of the force direction, the dataset includes examples that reflect real-world scenarios where tool compression may not be exactly normal to the surface of the breast. These variations are generated using FE simulations in FEBio [20], implemented through the GIBBON toolbox [21].

The dataset consists of $n = 1100$ deformed breasts with $m = 10859$ nodes, each node possessing x, y, z coordinates, and elasticity modulus E, which are illustrated in Fig. 3. It is divided into validation, testing, and training sets, with 10%, 10%, and 80% of the samples assigned to each set, respectively. Furthermore, 20 new samples were generated with mechanical properties (elasticity modulus, poison ratio) different from those in the dataset to establish accurate representations of deformed breasts containing tumours as the simulated ground truth. Utilizing this carefully crafted dataset allows for a detailed understanding and simulation of breast tissue dynamics and tracking its tissue motion. A leap in modelling fidelity based on this dataset and the proposed framework, which is discussed next, is anticipated to directly translate into enhanced accuracy in tumour localization and removal strategies.

III. The Proposed Framework

The proposed method's flowchart is depicted in Fig. 4 where the output of the FEM serves as input to the module, which approximates the true states and parameters by combining the complex model of deformed tissue observations with a theoretical model, utilizing ensemble-based methods to reduce computational costs. The module comprises two steps: first, utilizing a GNN-VAE network to compute distributions of latent variables associated with FEM output meshes, and second, employing the ES-MDA method to update prior distributions of latent variables using sophisticated measurements, resulting in the selection of the most probable combination of latent variables to construct an updated simulated mesh compatible with ground truth measurements. The following section provides a brief overview of GNN-VAE and then elaborates on the application of data assimilation using the ES-MDA.

A. Variational Auto-Encoder

Autoencoders are a subset of unsupervised neural networks widely applied for learning representations and reducing dimensions. They consist of an encoder, extracting low-dimensional features from high-dimensional input, and a decoder, reconstructing input data while minimizing errors [22]. Within the architecture of the VAE, $E_\theta(x)$ generates a latent-variable distribution (z) from $x \sim p(x)$ while $D_\theta(z)$ receives a random argument $z \sim p(z)$ and generates a sample from the learned distribution $\hat{x} \sim p_\theta(\mathrm{x} \mid \mathrm{z})$. In the proposed framework, (x) denotes deformed mesh output from FEM simulation, with training samples derived from patient-specific meshes. The generative model facilitates the easy generation of new meshes post-training. The goal is to update the prior distribution of latent variables through data assimilation from ground truth measurements, obtaining the posterior distribution and sampling latent variables with the highest probabilities.

B. Graph Neural Network

Within the architecture of the VAE, the networks $E_\theta(x)$ and $D_\theta(x)$ consist of layers that are designed to align with the data structures they process. While images and time series typically exist within Euclidean domains, tetrahedral meshes are situated in irregular and non-Euclidean domains, which can be effectively represented using graphs. Applying conventional 2D or 3D convolutional networks directly to mesh data faces challenges due to irregularities in local structures, such as varying vertex degrees and sampling densities. Graph neural networks are specifically designed to extract information from graph data structures [23].

In [23], an autoencoder utilizing spatially variable convolution kernels has been introduced, where each vertex possesses its own convolution kernel. A vertex-specific kernel is estimated based on a global kernel weight foundation. Throughout the training process, both the global kernel weight foundation and a sampling function for each kernel are acquired. The spatially varying convolution layer is an efficient method for capturing spatially varying content in irregular mesh connections. In this study, the spatially varying convolution and pooling layers are utilized to construct the encoder and decoder of the VAE.

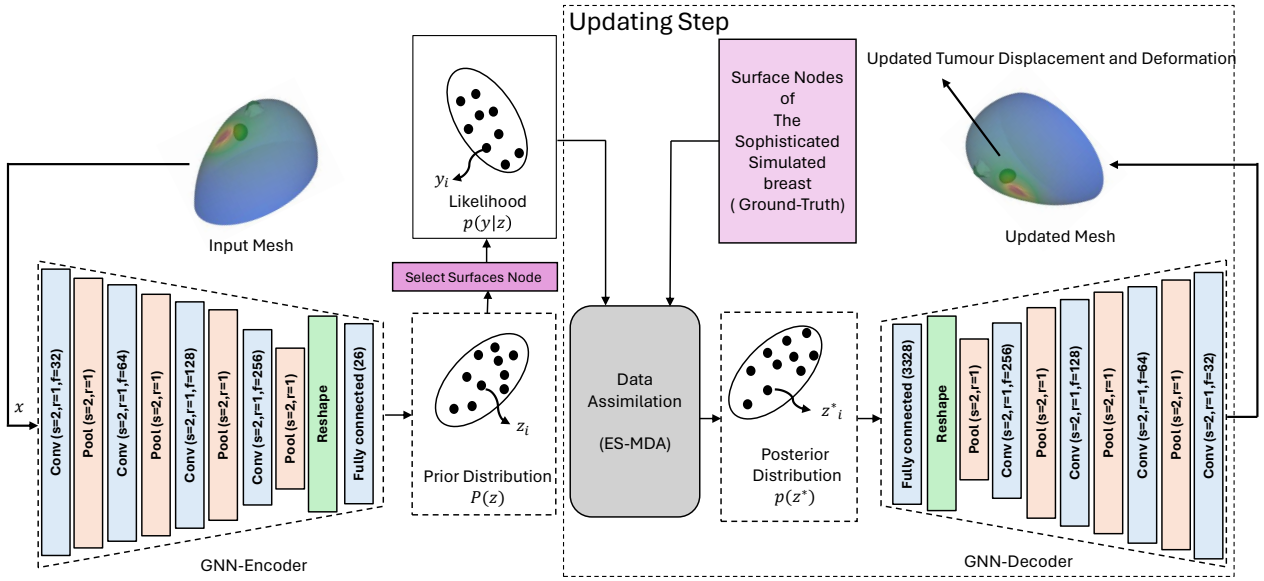

Fig. 4: The structure of the framework

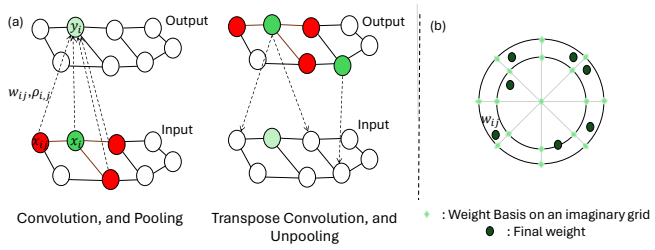

Fig. 5: (a) Graph convolutions and pooling layer on graph data with a radius of 1 and stride of 2; (b) global weight basis kernel introduced in [23].

1) Fully Convolution Graph Layer: In the convolutional layer, the input data $x \in \mathbb{R}^{V\times d}$ with V vertices and d dimension yields output data $y \in \mathbb{R}^{V\times d'}$ with d' dimension. A diagram illustrating the convolution operator is depicted in Fig. 5(a). The convolution operator for each vertex of a graph can be computed using:

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

Applying uniform weighting schemes to all vertices is not feasible because of the irregular distribution of vertices across a mesh and the varying connectivity between them. Each vertex needs flexibility to determine its convolution weight independently. In [23], a discrete convolution kernel is introduced with weights placed on a standard grid, referred to as the weight basis illustrated in Fig. 5(b). Vertices within a local region of the mesh are distributed across this grid. Weights at actual vertices can be obtained by sampling from the weight basis using different functions from vertex to vertex using the following equation:

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

2) Pooling Graph Layer: In a random graph, the distribution of vertices within the kernel radius can vary significantly, rendering methods like max or average pooling ineffective. To address this, a pooling layer is proposed in [23], which employs Monte Carlo sampling for feature aggregation as depicted in Fig. 5(a). A stride of 2 and a radius of 1 are utilized in this pooling layer. The resulting aggregated input nodes' output feature can be computed using:

$$\mathbf{y}_i = \sum_{j\in\mathcal{N}(i)} \rho'_{i,j} \mathbf{x}_{i,j} \tag{3}$$

where

$$\rho'_{i,j} = \frac{|\rho_{i,j}|}{\sum_{j=1}^{E_i} |\rho_{i,j}|}, \rho_{i,j} \in \mathbb{R} \tag{4}$$

C. Data assimilation with GNN-VAE

This paper presents a method for updating the output heterogeneous mesh of a breast model containing the tumour using only surface data without needing a complex temporal deep network. While Ensemble Kalman Filter (EnKF) is effective for sequential data assimilation, it requires a complex network, unlike Ensemble Smoother (ES), which computes a global update by assimilating all data simultaneously. However, the accuracy of one-step ES is constrained by the large ensemble update step sizes. To improve accuracy, the authors suggest using ES-MDA, which iteratively estimates unknown parameters using measurements at the current step [24]. This method involves assimilating the same data multiple times, enhancing accuracy. The unknown parameters are low-dimensional latent-space variables, and their relationship with observations is represented by a pre-trained decoder network. ES-MDA updates these estimates iteratively using available observations and the decoder network, accurately estimating latent-space variables compatible with observations.

The module follows a systematic approach. Firstly, at each time step, the output of the FEM simulation is input into the

module as illustrated in Fig. 4. Next, in the initialization step, prior ensembles are generated by sampling from a normal distribution in the latent space using the GNN-VAE encoder network. The number of iterations is then determined for subsequent steps. Subsequently, in the forecast step, each ensemble realization is passed through the GNN-VAE decoder network to generate an output mesh. From this mesh, the points on the surface are selected to predict the model's behaviour at specific measurement locations. Lastly, in the update step, the latent-space realizations are updated at each time step using a single set of measurements. To enable iterative data assimilation based on one measurement, the measurement vector is perturbed using a noise vector multiplied by an inflated covariance error matrix. The difference between the perturbed measurement vector and the ensemble predictions is calculated and weighted based on covariance matrices to enhance the accuracy of the ensemble prediction. Mathematically, the update rule can be expressed as shown in the following equation:

$$\begin{aligned}\mathbf{z}_i^{n+1} = \mathbf{z}_i^n + \mathbf{C}_{\mathbf{zy}}^n \left(\mathbf{C}_{\mathbf{yy}}^n + \alpha_n \mathbf{C}_{\mathbf{d}}\right)^{-1} \\ \left(\mathbf{d}_{\text{obs}} + \sqrt{\alpha_n}\mathbf{C}_{\mathbf{d}}^{1/2}\varepsilon_i^n - \mathbf{y}_i^n\right)\end{aligned} \tag{5}$$

In this equation, $\mathbf{C_d}$ represents a covariance matrix defined by the user, and ε_i^n denotes the observation error at iteration n, which follows a Gaussian distribution $\mathcal{N}(0, IN_d)$. Here, N_d represents the number of observations. The coefficient α_n adjusts the measurement error and covariance matrix at each iteration n. These coefficients are chosen in decreasing order to gradually reduce the impact of initial updates, particularly when there is a significant difference between predictions and observations in the early iterations. The coefficients α_n must also satisfy the condition $\sum_{n=1}^{N_a} \frac{1}{\alpha_n} = 1$, where N_a is the total number of iterations. The matrices $\mathbf{C}_{\mathbf{zy}}^n$ and $\mathbf{C}_{\mathbf{yy}}^n$ represent the cross-covariance matrix between latent-space variables and surface point predictions and the autocovariance matrix of surface point predictions, respectively. These matrices are computed from the ensemble at each iteration (n) using :

$$\begin{aligned}\mathbf{C}_{\mathbf{zy}}^n &= \frac{1}{N_e - 1}\sum_{i=1}^{N_e}\left(\mathbf{z}_i^n - \overline{\mathbf{z}}\right)\left(\mathbf{y}_i^n - \overline{\mathbf{y}}\right)^T \\ \mathbf{C}_{\mathbf{yy}}^n &= \frac{1}{N_e - 1}\sum_{i=1}^{N_e}\left(\mathbf{y}_i^n - \overline{\mathbf{y}}\right)\left(\mathbf{y}_i^n - \overline{\mathbf{y}}\right)^T\end{aligned} \tag{6}$$

In this context, N_e represents the total ensemble realizations. $\overline{\mathbf{z}}$ denotes the ensemble mean of the latent-space variables, and $\overline{\mathbf{y}}$ represents the ensemble mean of the surface point predictions. All the detailed steps are outlined in Fig. 6.

IV. Simulation and Results

In this section, we detail the simulation setup designed to evaluate the effectiveness of our proposed framework for breast tumour-tissue tracking. Each simulation is carefully structured to investigate the influence of key parameters, including the number of ensembles (N_e), update steps (N_a), and the measurement covariance matrix (C_d), on the precision of tumour localization. These parameters were selected based on their potential impact on model performance and computational efficiency, reflecting scenarios that closely mimic real-world surgical conditions. By systematically varying these parameters, we aim to identify optimal configurations that maximize accuracy while minimizing computational demands, which are crucial for surgical applications. For this reason, first, the GNN-VAE was trained using data acquired from FEM simulations detailed in Section II. The accuracy of the trained network on training data is 0.014 cm, and on testing data is 0.026 cm. For assessing the impact of parameters N_e, N_a, and C_d on the accuracy of the registrations, ten random meshes from the test dataset were chosen as initial meshes within the framework. Additionally, ten different random meshes from the ground-truth dataset were selected as target meshes, corresponding to each initial mesh.

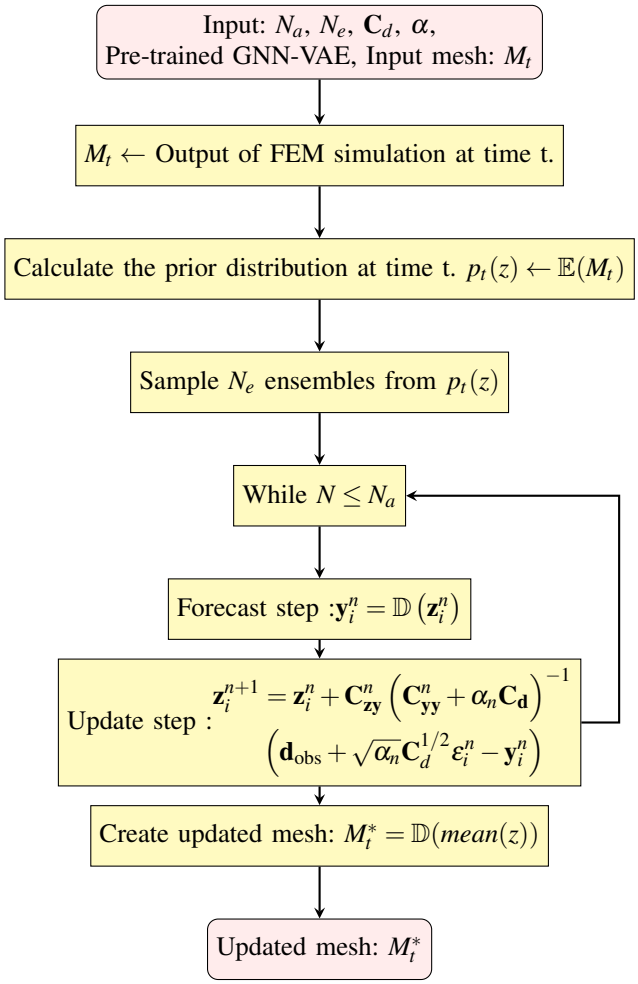

Fig. 6: The proposed method flowchart.

In Fig. 7, we maintain $N_a = 5$ constant and analyze the effects of varying N_e and C_d. Taking C_d down from 0.1 to 0.001 produces a lower average mean squared error.Fig. 7 also indicates that there is not a significant difference between $C_d = 0.001$ and $C_d = 0.01$. Furthermore, selecting $N_e = 10$ leads to poorer performance compared to $N_e = 60$. It can be concluded that increasing the N_e allows for a

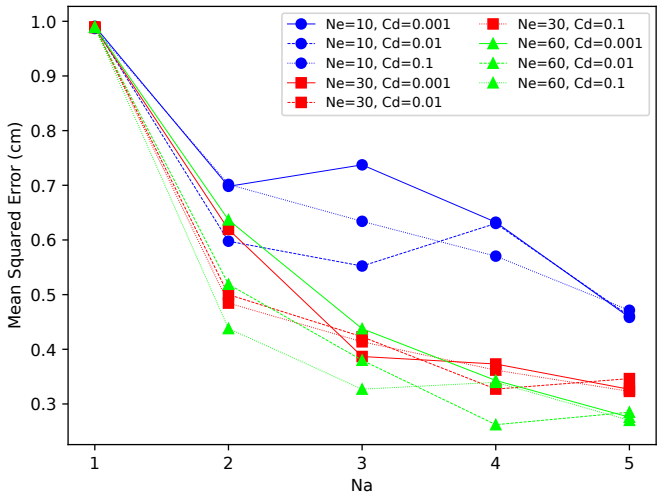

Fig. 7: An examination of how the numerical parameters C_d and N_e impact the MSE (cm).

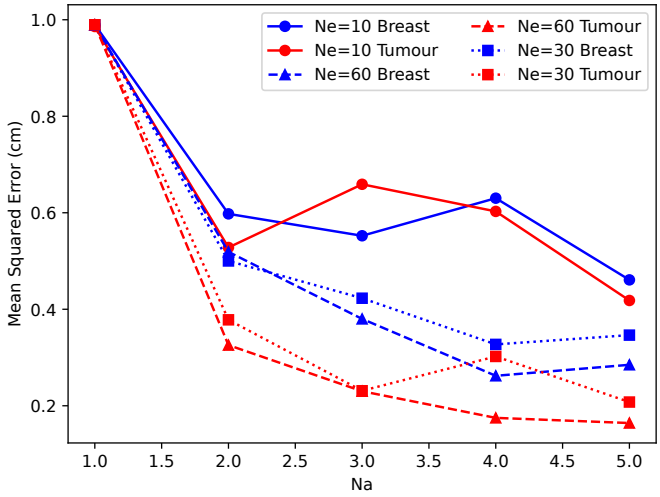

Fig. 8: The effect of Variation of the N_e and N_a while the $C_d = 0.01$ is constant on tumour localization accuracy.

smaller MSE. However, it is crucial to consider the associated increase in computational cost and time required for updates, especially in surgical applications. Fig. 7 illustrates that the results for $N_e = 30$ are comparable to those for $N_e = 60$. The simulations were conducted using a computational resource with 64 GB of RAM, 16 cores of CPU running at 3.40 GHz, and GPU Nvidia RTX 4070 with 12 GB VRAM. The MSE did not change significantly when the number of ensembles increased beyond 60. However, for $N_e > 60$, the time required for updates increases, which may not be beneficial in practical applications.

The main application of this paper was simulating precise tumour tracking in a lumpectomy. For this reason, the accuracy of tumour localization was investigated by varying the parameters N_a and N_e while keeping $C_d = 0.01$, as shown in Fig. 8. Interestingly, we observed a consistent pattern in the tumour and breast tissue; however, after updating the tumour position, we achieved higher accuracy than the breast tissue.

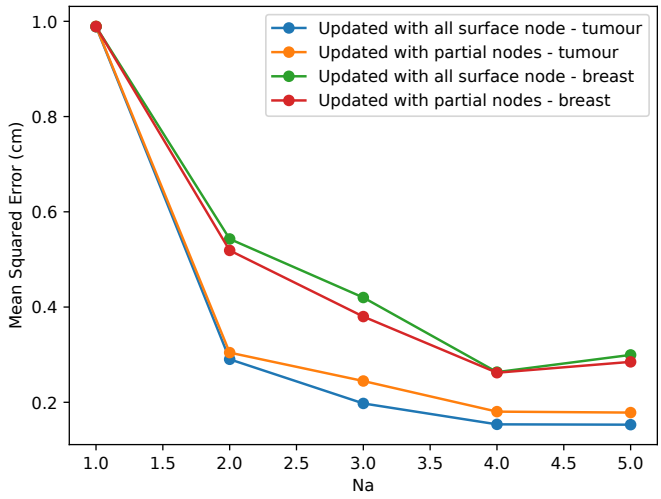

Fig. 9: The effect of partial number of surface nodes on the MSE(cm), $N_a = 1:5$ while the $N_e = 60$, $C_d = 0.01$ is constant on tumour localization accuracy.

This improvement can be attributed to the stiffness of the tumour, which results in lower displacement. Consequently, we observed a lower MSE after the update, indicating a more accurate tumour localization.

One critical parameter in the updating step is the ground truth surface nodes, which play a crucial role in assessing the effectiveness of the proposed method. In this paper, these data are sourced from a more sophisticated simulation to represent the ground truth. The information is then used to update all nodes within the tissue, as the framework's main advantage lies in updating the inside of the breast using surface information. During real lumpectomy, tools like ultrasound probes may obscure some surface nodes, making fewer nodes available for updating than the total surface nodes. In addition, real tools for tracking the surface nodes have the limitations of precisely detecting all the surface nodes. Two updating scenarios were considered. In the first scenario, tumour deformation and position were updated using all nodes on the surface of the ground truth breast, and the MSE was recorded for the updated tumour position and deformation. In the second scenario, some of the surface nodes were available for updating due to the inability to track all of them in real surgery.

Fig. 9 compares the MSE for two scenarios: updating using all surface nodes and updating using partial surface nodes. The analysis shows minimal differences in MSE between these scenarios for both tumour tracking and overall tissue changes. This suggests that even if some points on the tissue surface are missed during the procedure, whether due to human error or limitations of surgical tools, the updating process remains effective. Moreover, utilizing a partial number of surface nodes for updating can effectively reduce computational costs and time requirements, making it a viable option.

The results presented above not only validate the effectiveness of our proposed framework in achieving high

accuracy in tumour tracking during simulated lumpectomies but also underscore its superiority over existing methods. By analyzing the impact of various parameters on the MSE, we demonstrate a significant enhancement in precision. This represents a step forward in the application of non-invasive, precise tumour-tracking technologies, directly contributing to the improvement of lumpectomy procedures and patient care.

V. Conclusions

This paper presents an innovative approach to tracking deformable tumours within breasts of varying mechanical properties during simulated lumpectomy procedures. We introduce an online residual model that integrates physics-based tissue simulations using FEM with measurements of deformed tumours and breasts, including tumour displacement. Our novel framework incorporates GNN-VAE and ES-MDA to align simulation outputs with more sophisticated simulation data. To address the challenge of real-time updates to high-dimensional mesh models for tissue deformation, we utilize generative auto-encoder networks to learn simulation-data distributions and ES-MDA to update these distributions with the ground truth surface nodes. The results indicate that the MSE achieved under varying parameters underscores the method's precision. Additionally, the framework demonstrates robustness in tumour tracking during a simulated lumpectomy, even when utilizing a partial number of surface nodes for updating. This suggests that the approach is resilient to missing surface data points, which are common in surgical settings due to the limitations of current tracking tools. It is important to emphasize that our research simulates tumour tracking to evaluate the proposed method's capability for real surgery. However, real breast tissue would replace the simulated ground truth used in this paper in actual surgical scenarios. Future work will involve integrating real deformed breasts with deformable tumours as the ground truth, further validating the effectiveness of the approach in real clinical scenarios.

References

[1] I. Hoxha, F. Sadiku, L. Hoxha, M. Nasim, M. A. C. Buteau, K. Grezda, and M. D. Chamberlin, "Breast cancer and lifestyle factors: Umbrella review," *Hematology/Oncology Clinics*, vol. 38, no. 1, pp. 137–170, 2024.

[2] S. Zargarzadeh, A. Sieben, E. Wiebe, L. Peiris, and M. Tavakoli, "Augmented reality-based tumor localization and visualization for robot-assisted breast surgeries," in *2024 IEEE 4th International Conference on Human-Machine Systems (ICHMS)*. IEEE, 2024, pp. 1–6.

[3] J. H. Hipwell, V. Vavourakis, L. Han, T. Mertzanidou, B. Eiben, and D. J. Hawkes, "A review of biomechanically informed breast image registration," *Physics in Medicine & Biology*, vol. 61, no. 2, p. R1, 2016.

[4] M. M. Kapoor, M. M. Patel, and M. E. Scoggins, "The wire and beyond: recent advances in breast imaging preoperative needle localization," *Radiographics*, vol. 39, no. 7, pp. 1886–1906, 2019.

[5] L. Gabrielova, I. Selingerova, J. Zatecky, O. Zapletal, P. Burkon, M. Holanek, and O. Coufal, "Comparison of 3 different systems for non-wire localization of lesions in breast cancer surgery," *Clinical Breast Cancer*, 2023.

[6] S. Shirazi, H. Hajiesmaeili, M. Khosla, S. Taj, T. Sircar, and R. Vidya, "Comparison of wire and non-wire localisation techniques in breast cancer surgery: a review of the literature with pooled analysis," *Medicina*, vol. 59, no. 7, p. 1297, 2023.

[7] F. Martínez-Martínez, M. J. Rupérez-Moreno, M. Martínez-Sober, J. A. Solves-Llorens, D. Lorente, A. Serrano-López, S. Martínez-Sanchis, C. Monserrat, and J. D. Martín-Guerrero, "A finite element-based machine learning approach for modeling the mechanical behavior of the breast tissues under compression in real-time," *Computers in biology and medicine*, vol. 90, pp. 116–124, 2017.

[8] M. Afshar, T. Meyer, R. S. Sloboda, S. Hussain, N. Usmani, and M. Tavakoli, "Registration of deformed tissue: A gnn-vae approach with data assimilation for sim-to-real transfer," *IEEE/ASME Transactions on Mechatronics*, 2023.

[9] 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: Solid and Fluid Mechanics for the Benefit of Patients 22*. Springer, 2020, pp. 33–45.

[10] 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.

[11] 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.

[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] 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.

[14] M. Alakuijala, G. Dulac-Arnold, J. Mairal, J. Ponce, and C. Schmid, "Residual reinforcement learning from demonstrations," *arXiv preprint arXiv:2106.08050*, 2021.

[15] 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.

[16] 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.

[17] 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*, vol. 27, no. 4, pp. 2194–2203, 2022.

[18] S. K. Warfield, M. Ferrant, X. Gallez, A. Nabavi, F. A. Jolesz, and R. Kikinis, "Real-time biomechanical simulation of volumetric brain deformation for image guided neurosurgery," in *SC'00: Proceedings of the 2000 ACM/IEEE Conference on Supercomputing*. IEEE, 2000, pp. 23–23.

[19] N. G. Ramiao, P. S. Martins, R. Rynkevic, A. A. Fernandes, M. Barroso, and D. C. Santos, "Biomechanical properties of breast tissue, a state-of-the-art review," *Biomechanics and modeling in mechanobiology*, vol. 15, pp. 1307–1323, 2016.

[20] S. A. Maas, B. J. Ellis, G. A. Ateshian, and J. A. Weiss, "Febio: finite elements for biomechanics," 2012.

[21] K. M. Moerman, "Gibbon: the geometry and image-based bioengineering add-on," *Journal of Open Source Software*, vol. 3, no. 22, p. 506, 2018.

[22] Y. Wang, H. Yao, and S. Zhao, "Auto-encoder based dimensionality reduction," *Neurocomputing*, vol. 184, pp. 232–242, 2016.

[23] 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.

[24] A. A. Emerick and A. C. Reynolds, "Ensemble smoother with multiple data assimilation," *Computers & Geosciences*, vol. 55, pp. 3–15, 2013.