Real-time Prediction of Breast Cancer Site using Graph Neural Networks

Kyunghyun Lee

The Graduate School
Yonsei University
Department of Computational Science and Engineering

Real-time Prediction of Breast Cancer Site using Graph Neural Networks

A Master's Thesis Submitted
to the Department of Computational Science and Engineering
and the Graduate School of Yonsei University
in partial fulfillment of the
requirements for the degree of
Master of Mathematics

Kyunghyun Lee

December 2024

This certifies that the Master's Thesis of Kyunghyun Lee is approved

Thesis Supervisor	Kyungho Yoon
Thesis Committee Member	Wonyong Shin
Thesis Committee Member	Eunjung Lee

The Graduate School
Yonsei University
December 2024

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABSTRACT

Real-time Prediction of Breast Cancer Site using Graph Neural Network

Breast cancer is the most common cause of cancer-related fatalities in women, but early diagnosis significantly improves treatment outcomes. Breast MRI plays a critical role in early detection, especially for lesions that mammography and ultrasound may miss. However, confirming malignancy requires a biopsy, a procedure that involves extracting tissue samples for pathological analysis. Direct MRI-guided biopsies are highly accurate but costly and time-consuming. To address these limitations, an indirect MRI-guided breast biopsy system has been proposed recently, which leverages MRI-guided navigation to perform biopsies outside the MRI room. However, traditional navigation systems have limitations, as they cannot account for tissue deformation in real time. To solve this issue, we developed a real-time 3D deformable model using Graph Neural Network (GNN). Our model predicts breast tissue deformations based on surface displacements, bypassing the need for force measurements. Datasets were created using finite element simulations on 3D models generated from breast phantoms and real patient MRI data, incorporating varying material properties and force conditions. The input features consist of the displacement of surface nodes, while the output features consist of the displacement of all nodes. Our GNN model, designed with a specialized graph structure to represent physical properties, was trained on these datasets to predict output features from input features. As a result, with our best model, we achieved average RMSE of 0.888 mm, average cancer RMSE of 0.089 mm, and dice score of 0.928 for phantom based dataset. Additionally, the same model showed prominent performance in terms of RMSE and dice score on real patient data.

Key words : **GNN, Biopsy, Breast Cancer, MRI**

1. Introduction

1.1. Research background

Breast cancer is the most prevalent cancer and fatal cause of cancer-related deaths among women [1]. Nevertheless, early diagnosis holds the potential to expand treatment options and improve survival rates [2].

Among many diagnosis tools for breast cancer, breast MRI is crucial for early detection of breast cancer since over half of lesions are exclusively detectable through MRI, capturing early-stage breast cancers evading detection by other tests such as mammography and ultrasound [3]. However, MRI alone cannot assure that the lesion is cancerous. Biopsy is needed to confirm that the lesion is cancerous.

Biopsy is the most definitive method for diagnosing cancer, as it involves taking and examining tissue samples from areas of concern. Among the various biopsy techniques, core needle biopsy is widely used [4]. Direct MRI-guided breast biopsy, in particular, uses MRI to identify and sample abnormalities in the breast, especially those that are challenging to detect with mammography or ultrasound [5]. Although this approach provides excellent accuracy, especially for complex or hard-to-detect lesions, it can be both expensive and time-consuming.

Thus, indirect MRI-guided breast biopsy is proposed recently. Instead of performing biopsy in MRI room, MRI-guided navigation system that can move MRI space and real space interchangeably can be used, allowing the biopsy to be performed outside the MRI room. MRI-guided navigation system is already used in medical field. For instance, [6] have investigated the feasibility of MRI-guided techniques for demarcating breast cancer in preparation for breast-conserving surgery. They explored strategies to enhance the accuracy and effectiveness of MRI-guided navigation, particularly in the context of delineating breast tumors in supine patient orientation. However, there is a limitation to the MRI-guided navigation system.

For biopsies, the navigation system must show real-time deformation because any external interference (such as a needle or finger) can cause the cancer to move. Unfortunately, MRI provides only still images, so it cannot display deformation. Therefore, traditional MRI-guided navigation systems are not suitable for core needle biopsies. Thus, for indirect MRI-guided breast biopsy, real-time 3D deformable model is needed, because current 3D deformable model takes several tens of minutes.

Therefore, our objective is 3D deformable model that can predict breast cancer sites. The model must have short inference time for real-time prediction. There have been some efforts to use machine learning techniques to shorten the inference time, particularly by utilizing geometric information [7, 8]. Furthermore, there have been several attempts to develop surrogate models for FEM using GNNs [9-13], among which PhysGNN [13] stands out as a notable example. The authors employed Graph Neural Network (GNN) to predict brain tissue deformations under applied forces. However, measuring forces poses significant challenges during biopsies. Therefore, our aim is to develop a GNN that can predict breast cancer sites using surface deformation data, which is accessible.

1.2. Method Overview

The methodologies employed in this study are detailed as follows. First, a finite element model is generated using MRI data from either a breast phantom or a real patient's breast. Next, various loading conditions are simulated through finite element analysis to create a training dataset. The input data consist of displacements of surface nodes, while the output represents the displacements of all nodes. This dataset is then used to train a GNN.

Four types of datasets were created using breast phantoms:

1. Single force location with linear uniform material. (Dataset 1)
2. Multi-force locations with linear uniform material. (Dataset 2)
3. Multi-force locations with nonlinear uniform material. (Dataset 3)
4. Multi-force locations with nonlinear separated material. (Dataset 4)

These datasets progressively incorporate more realistic conditions by varying properties such as force location and material properties.

The same methodology was applied to real patient data (Dataset 5-7), specifically using datasets with multi-force locations and nonlinear separated material. For validation, three datasets from real patient data were utilized to assess the model developed with breast phantom datasets.

This workflow highlights the capability of the GNN to predict displacement of all nodes from displacement of surface nodes, enabling the development of a real-time deformable model based on the provided data.

2. Methods

The methods section is divided into two subsections: Data Acquisition and GNN Training. The Data Acquisition subsection is further divided into three procedures: Pre-Processing, Finite Element Simulation, and Dataset Generation. Fig 1 shows the whole process of the method.

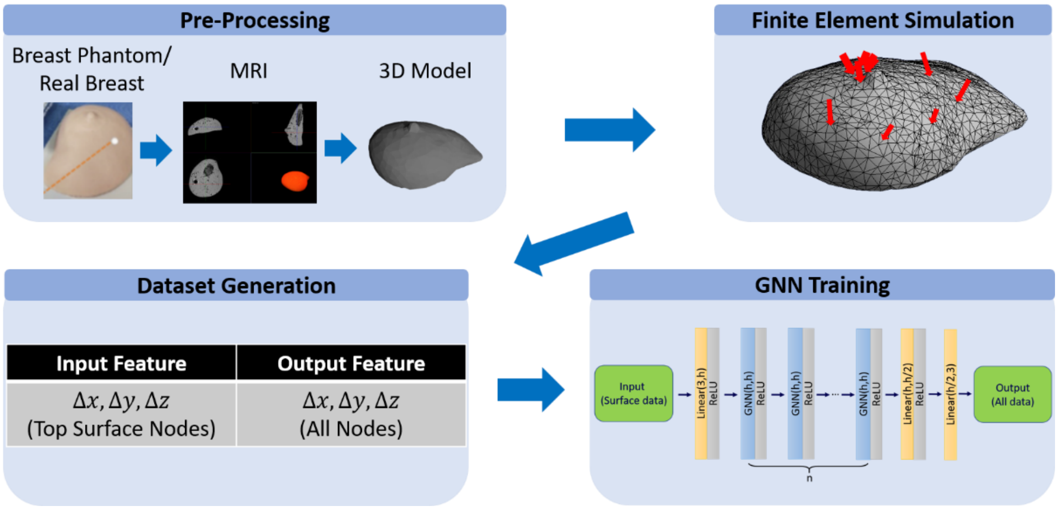
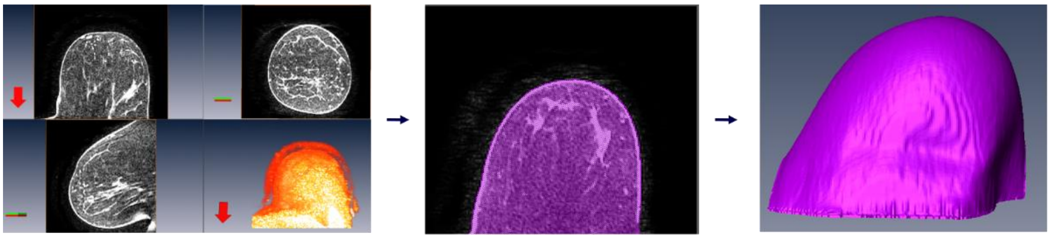


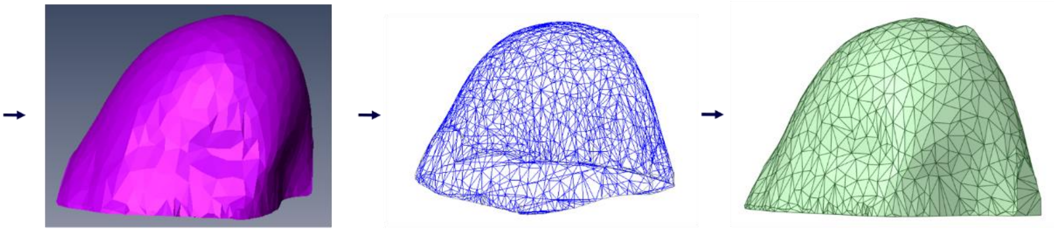
Fig 1. Whole Process of the Method. The methods section is divided into two subsections: Data Acquisition and GNN Training. The Data Acquisition subsection is further divided into three procedures: pre-processing, finite element Simulation, and dataset generation.

2.1. Data Acquisition

Pre-processing We first acquired MRI data from breast phantoms or real patients. Using a 3D image visualization and analysis tool, we processed the MRI data to generate a surface model of the breast. This involved segmentation using the threshold and brush tools to isolate the breast region and cancer region and create a detailed surface representation. The resulting surface was then refined and simplified, making it suitable for further processing. The simplified surface was exported as an OBJ file, which was then imported into ANSYS SpaceClaim to create a 3D breast model. The entire modeling workflow is illustrated in Fig. 2.



(a) MRI data (b) Segmentation (c) SurfaceGen



(d) Simplifier (e) Obj file (f) SpaceClaim

Fig 2. Process of 3D modeling. First, MRI of breast phantom or breast of real patient is taken. Then segmentation is proceeded using the threshold and brush tools to isolate the breast region and cancer region obtain. The resulting surface was then refined and simplified using the simplifier tool in AMIRA obj file from that. The simplified surface was exported as an OBJ file, which was then imported into ANSYS SpaceClaim to create a 3D breast model.

Finite Element Analysis For Dataset 1-2, linear finite element analysis was applied. Continuum mechanics based local equilibrium equations govern the tissue deformation in linear analysis [14], as indicated by

$$\tau_{ij,j} + f_i^B = 0 \text{ in } \Omega, \tag{1}$$

where $\tau_{ij,j}$ denotes the derivate of the ij component of the stress tensor with respect to j, f_i^Brepresents the i-directional component of the applied body force, and Ω indicates the volume region under analysis. From Eq. 1, principle of virtual work (PVW) is derived.

$$\int_V \tau_{ij,j}\delta\epsilon_{ij}dV = \int_V f_i^B\delta u_i dV + \int_{S_f} f_i^S\delta\epsilon_{ij}dS \tag{2}$$

V is volume, τ is stress tensor, ϵ is strain tensor, f^B is body force, S_f is surface force, δu is virtual strain and $\delta\epsilon$ is virtual displacement.

On the other hand, for Dataset 3-7, nonlinear finite element analysis was applied. Continuum mechanics based total Lagrangian formulation govern the tissue deformation in nonlinear analysis [14], as indicated by

$$\int_{{}^0V} \delta_0 e_{ij\,0}C_{ijrs\,0}e_{rs}\, d^0V + \int_{{}^0V} {}_0^tS_{ij}\,\delta_0\eta_{ij}\, d^0V = {}^{t+\Delta t}R - \int_{{}^0V} {}_0^tS_{ij}\,\delta_0 e_{ij}\, d^0V$$

Where ${}^0V$ is volume at time 0, $\delta_0 e_{ij}$ is virtual incremental Green Lagrange (GL) strain (linear), $\delta_0\eta_{ij}$ is virtual incremental GL strain (nonlinear), ${}_0^tS_{ij}$ is stress measured from the initial configuration, and ${}^{t+\Delta t}R$ is external virtual work at time $t+\Delta t$.

This further can be changed into ${}_0^tKU = {}^{t+\Delta t}R - {}_0^tF$ form, where ${}_0^tK$ is tangential stiffness matrix, U is incremental nodal displacement vector, ${}^{t+\Delta t}R$ is external nodal force vector at time $t+\Delta t$, and ${}_0^tF$ is internal force vector.

After applying appropriate mechanics for each dataset, we separated the elements and apply PVW or total Lagrangian formulation to each element individually before combining them together. These entire processes were automatically computed in ANSYS Mechanical. The automation provided by ANSYS Mechanical allows for efficient and accurate calculations.

Finite Element Simulation The simulation was performed using finite element analysis under conditions specific to each dataset. Random forces were applied to nodes near one of the fixed reference nodes. Specifically, the forces were applied to nodes within a sphere centered on the reference node, hereafter referred to as force sphere, with a radius randomly chosen between specific range. The bottom surface nodes are considered to be fixed and used as the boundary condition. The direction of force is random, but y-direction force is negative, where y-direction is height direction. Magnitude of force is random also, in particular range. Using ANSYS Mechanical, we calculated nodal displacement of multiple simulations for each dataset.

Dataset Generation Datasets were created based on the results of the ANSYS analysis. The analysis provided the displacement data for all nodes, but only the displacements of the top surface nodes were used as input, while the displacements of all nodes were used as output. Thus, the goal with this dataset is to develop a GNN model capable of predicting the nodal displacements of all nodes using the displacements of only the top surface nodes. Since a GNN requires node features for all nodes, we assigned zero values to the input features of nodes that were not on the top surface. Additionally, fixed nodes were excluded from both the input and output data. Table 1 organizes the input and output information.

Dataset 1–4 are phantom-based, while Dataset 5–7 are real patient-based. As the process progresses through Dataset 1–4, the model becomes increasingly complex and realistic. From Dataset 1 to Dataset 2, variations are introduced in the reference node, which determines the force location. In Dataset 3, the material property is changed to nonlinear. Finally, in Dataset 4, cancer is represented as a mass, requiring the use of two separate material properties. The conditions established in Dataset 4 are also applied to Dataset 5–7, with the only difference being that the MRI data in Dataset 5–7 is based on real patients rather than phantoms. Table 2 organizes the property of each dataset.

Table 1. Input and Output Information. The displacement of the top surface nodes is input, and the displacement of all nodes is output. zero value is assigned for nodes other than top surface nodes in input.

Input	**Top Surface Nodes**	Nodal Displacement $(\Delta x, \Delta y, \Delta z)$
	Other Nodes	0 (0, 0, 0)
Output	**All Nodes**	Nodal Displacement $(\Delta x, \Delta y, \Delta z)$

Table 2. Property of Each Dataset. Dataset 1–4 are phantom-based, while Dataset 5–7 are real patient-based. As the process progresses through Dataset 1–4, the model becomes increasingly complex and realistic. From Dataset 1 to Dataset 2, variations are introduced in the reference node. In Dataset 3, the material property is changed to nonlinear. In Dataset 4, cancer is represented as a mass. The conditions established in Dataset 4 are also applied to Dataset 5–7, with the only difference being that the MRI data in Dataset 5–7 is based on real patients rather than phantoms.

Dataset	Reference node	Material property	Cancer	MRI	Dataset size
Dataset 1	1	Linear	Node	Phantom	1000
Dataset 2	10	Linear	Node	Phantom	1000
Dataset 3	10	Nonlinear	Node	Phantom	1000
Dataset 4	10	Nonlinear	Mass	Phantom	3000
Dataset 5-7	10	Nonlinear	Mass	Real patient	3000

2.1.1. Breast Phantom Based Dataset

Dataset 1

The mesh is set with a target quality of 0.05, resulting in an average surface area of 2.6473×10^{-5} m^2 and a minimum edge length of 2.2894×10^{-3} m. For constructing the FE model, 10-node tetrahedral quadratic elements (known as Tet10 in ANSYS) are assigned to the meshes, totaling 7,598 elements and 13,311 nodes. Additionally, we assumed linear material properties in the finite element model. In this model, we set Young's modulus to 6,900Pa and Poisson's Ratio to 0.495.

Reference nodes for force location is fixed one node. The radius of force sphere ranges from 0.02m to 0.032m. The force magnitude ranges from 12N to 24N, and this force is evenly distributed across the nodes in the force sphere. Which means that the force applied to each node is between 12N/M and 24N/M, where M represents the total number of nodes in force sphere. The dataset consists of 1,000 simulation results, with 800 used for the training set, 100 for the validation set, and 100 for the test set.

Dataset 2

Dataset 2 shares all properties with Dataset 1, except for the reference nodes. In Dataset 2, reference nodes for force locations are fixed to 10 specific points, with 5 located near the cancer and the other 5 positioned farther away. This dataset includes 1,000 simulation results, with 100 simulations per reference node. Similar with Dataset 1, the data is split into 800 simulations for the training set, 100 for the validation set, and 100 for the test set.

Dataset 3

Dataset 3 shares all properties with Dataset 2, except for the material property. We assumed nonlinear material properties in the finite element model. Among the various models available to represent nonlinear material properties, we opted for the Mooney-Rivlin 2-Parameter model. In this model, we set C_{10} to 2000 Pa and C_{01} to 1333 Pa for fat tissue [15]. This dataset includes 1,000 simulation results, with 100 simulations per reference node. Similar with Dataset 1, 2, the data is split into 800 simulations for the training set, 100 for the validation set, and 100 for the test set.

Dataset 4

The mesh is set with a target quality of 0.05, resulting in an average surface area of $2.2987 \times 10^{-5}\ \mathrm{m}^2$ and a minimum edge length of $1.6802 \times 10^{-3}\ \mathrm{m}$. For constructing the FE model, Tet10 elements are assigned to the meshes, totaling 9,219 elements and 15,864 nodes. Additionally, we assumed nonlinear separated material properties in the finite element model. There are 2 materials in the model: fat and cancer. We opted for the Mooney-Rivlin 2-Parameter model in Dataset 4 also. In this model, we set C_{10} to 2000 Pa and C_{01} to 1333 Pa for fat tissue, and C_{10} to 6667 Pa and C_{01} to 10000 Pa for cancer tissue [15].

Reference nodes for force locations are fixed 10 locations, where 5 locations are near to the cancer and other 5 locations are far from the cancer. The radius of force sphere ranges from 0.02m to 0.032m. The force magnitude ranges from 120N to 240N, and this force is evenly distributed across the nodes in the force sphere. Which means that the force applied to each node is between 120N/M and 240N/M, where M represents the total number of nodes in force sphere. The dataset consists of 3,000 simulation results, with 2,400 used for the training set, 300 for the validation set, and 300 for the test set.

2.1.2. Real Patient Based Dataset

Basically, most of the properties of each dataset are the same as those of Dataset 4. Therefore, only the differences from Dataset 4 are described in the following parts.

Dataset 5

The mesh was generated with a target quality of 0.05, yielding an average surface area of $4.0017 \times 10^{-5}\ \mathrm{m}^2$ and a minimum edge length of 7.0702×10^{-4} m. It consists of 9,744 elements and 16,985 nodes.

Dataset 6

The mesh was generated with a target quality of 0.05, yielding an average surface area of $2.4443 \times 10^{-5}\ \mathrm{m}^2$ and a minimum edge length of 9.4996×10^{-4} m. It consists of 8,562 elements and 15,064 nodes.

Dataset 7

The mesh was generated with a target quality of 0.05, yielding an average surface area of $2.9163 \times 10^{-5}\ \mathrm{m}^2$ and a minimum edge length of 1.0066×10^{-3} m. It consists of 9,810 elements and 16,935 nodes.

2.2. GNN Training

2.2.1 Graph Structure

GNN [16] is a neural network architecture specifically designed to handle structured data represented in the form of graphs. It is usually used to task such as node classification, graph classification. Graph can be expressed as $\mathrm{G} = (\mathrm{V}, \mathrm{E})$, where V is the set of nodes and E is the set of edges. Additionally, each node has node features that can express properties of the node. Fig 3 illustrates the definition of Graph.

In GNN, message passing is a fundamental process, where each node collects, or aggregates, information from its neighboring nodes using edge information to update its own features. Fig 4 illustrates message passing of GNN.

Therefore, we have to define node condition and edge condition to construct the graph and utilize message passing. In this model, we made a graph satisfying the following conditions to represent the physical structure.

1. Nodes of Graph are the nodes of Finite Element Analysis
2. For Nodes A, B, $|\mathrm{r}_A - r_B| < d \Leftrightarrow A, B$ are connected, where r_A means the position of node A

d was set to 0.005m for Dataset 1-3, 0.003m for Dataset 4-7. Fig 5 shows the above conditions.

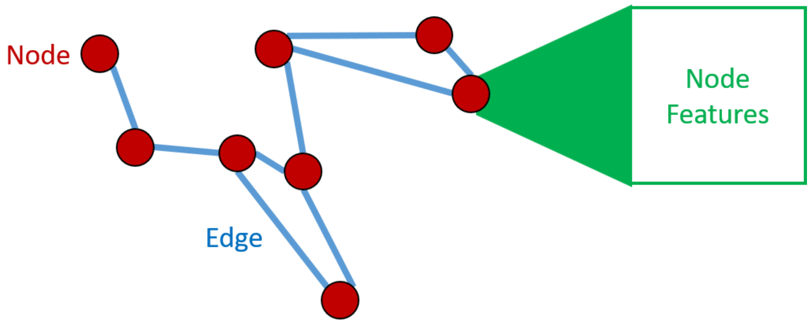


Fig 3. Definition of Graph. Graph can be expressed as $\mathrm{G} = (\mathrm{V}, \mathrm{E})$, where V, is the set of nodes and E is the set of edges. Additionally, each node has node features that can express properties of the node.

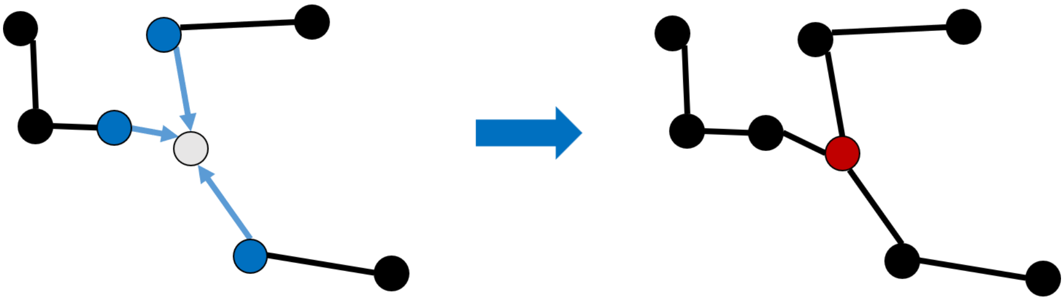
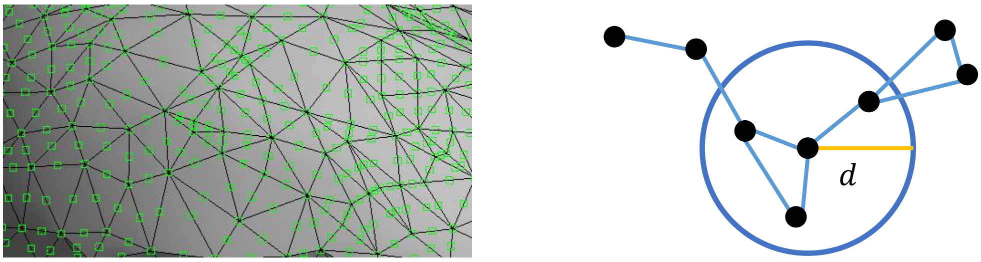


Fig 4. Message Passing of GNN. In GNN, message passing is a key mechanism, where each node aggregates information from its neighboring nodes using edge information to update its own features.



(a) Node Condition

(b) Edge Condition

Fig 5. Graph Structure. (a) Nodes of graph are the nodes of Finite Element Analysis. (b) For Nodes $A, B,\ |\mathrm{r}_A - r_B| < d \Leftrightarrow A, B$ are connected, where r_A means the position of node A.

Actually, edge condition could simply be defined as nodes within the same element being connected to each other. However, doing so might lead to excessive edges, resulting in longer computation time and potentially exacerbating oversmoothing. Therefore, the decision was made to determine edge condition using Euclidean distance.

In addition, we applied random edge addition. This process involves creating edges that connect cancer surface nodes to breast surface nodes. In Dataset 1-3, where there is only one cancer node, multiple edges were added between the cancer node and randomly selected breast surface nodes. For Dataset 4-7, which contain multiple cancer nodes, edges were randomly added between individual cancer nodes and breast surface nodes. These additional edges facilitate effective message passing between cancer and surface nodes, enabling a reduction in the number of GNN layers required while enhancing the prediction performance for cancer nodes. Without this random edge addition, the message passing distance is limited to d multiplied by the number of GNN layers, restricting the model's capacity to learn and propagate information effectively. Fig 6 illustrates random edge addition process.

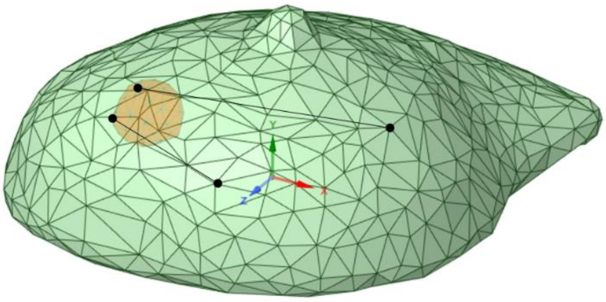


Fig 6. Random Edge Addition. Edges were added that connect one cancer surface node to one breast surface node randomly. These connections enable effective message passing between the cancer and surface nodes, allowing us to reduce the number of GNN layers needed while enhancing the prediction performance of cancer nodes.

Furthermore, for Dataset 4-7, to improve the accuracy of predicting the displacement of cancer nodes, a training mask was applied to the cancer nodes. This approach helps the model to focus on the cancer nodes during training, which results in more precise predictions for these critical nodes. While this may lead to less accurate predictions for the other nodes, the trade-off is a notable enhancement in prediction accuracy for the cancer nodes, which is the primary focus of the model.

2.2.2 GNN Model

Several layers of GNN and 3 linear layers are in our GNN model. For activation function, ReLU is used. For loss function, MSELoss is used, and for optimizer, AdamW is used. Learning rate is 1e-3. Total number of epochs are 1000 and best model which has the minimum validation loss after 500 epochs is saved. The number of GNN layers (n), the number of hidden channels (h), other hyperparameters and aggregation function can be adjusted based on the characteristics of the dataset. Fig 7 shows the structure of the GNN model.

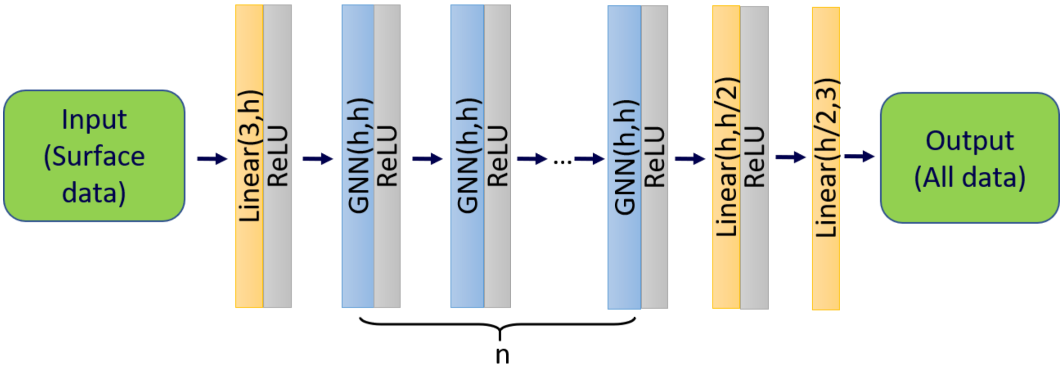


Fig 7. The Structure of GNN Model. Several layers of GNN and 3 linear layers are in our GNN model. For activation function, ReLU is used. For loss function, MSELoss is used, and for optimizer, AdamW is used. The number of GNN layers (n), the number of hidden channels (h), and other hyperparameters can be adjusted based on the characteristics of the dataset.

2.2.3 Computation Setting

Above methods are executed in this computer setting:

CPU: AMD Ryzen 9 5950X 16-Core Processor, 3401Mhz

RAM: 128GB

GPU: NVIDIA GeForce RTX 4080 (VRAM: 16GB)

3. Result

The main evaluation methods for this study are Average RMSE (Rooted Mean Square Error) of displacement and average cancer node error. RMSE is defined as follows:

$$\text{RMSE} = \frac{\sqrt{\Sigma(\delta_{GT} - \delta_{PD})^2}}{N}$$

where δ_{GT} is nodal displacement (ground truth simulated by ANSYS), δ_{PD} is nodal displacement (prediction of model), and N is the number of nodes. Therefore, we calculated the RMSE for each simulation in test dataset and then averaged RMSE values.

Evaluating this study using only RMSE is insufficient because most of the nodal displacements are zero, and this cannot directly show the prediction performance of cancer. Therefore, we also wanted to evaluate the displacement error of cancer nodes for each simulation. For datasets that used separated material, we calculated RMSE of cancer nodes. This is considered to be important evaluation method, because the object of this study is prediction of cancer site. On the other hand, for other dataset that cancer nodes are not specified, we chose one specific cancer center target node and calculated the average magnitude of displacement errors specifically for this target node in the test dataset.

However, RMSE itself cannot represent relative error. Thus, we calculated RMSD (Rooted Mean Square Displacement) of whole nodes and cancer nodes to evaluate relative accuracy of RMSE. RMSD is defined as follows:

$$\text{RMSD} = \frac{\sqrt{\Sigma(\delta_{GT})^2}}{N}$$

For datasets with uniform material properties, we calculated average cancer node displacement instead of cancer RMSD, because there is only one cancer node.

For datasets with separated material properties, we also evaluated the model's performance using the dice score, focusing specifically on the cancer mass. The dice score quantifies the overlap between the ground truth and predicted regions, providing insight into how accurately the model captures the shape and location of the cancer mass. This evaluation requires voxelizing the cancer mass with a resolution of 1 mm. The voxelization process starts with discretizing the coordinates of nodes by dividing each coordinate by the voxel size and applying the floor function, as given by:

$$w_{x_i} = \left\lfloor \frac{v_{x_i}}{s} \right\rfloor$$

where x_i means x, y, z (i =1, 2, 3, respectively), w_{x_i} is the voxelized coordinate, v_{x_i} is the position component, and s is the voxel size, set to 1mm. The resulting voxelized coordinates were stored as sets for both the ground truth and prediction.

Once voxelized, the intersection and union of the ground truth and predicted voxels are computed. The dice score is then defined as follows:

$$\text{Dice Score} = \frac{2 \times n(GT \cap PD)}{n(GT) + n(PD)}$$

where GT is set of voxels of ground truth and PD is set of voxels of prediction.

The evaluation methods used for each dataset is summarized in Table 3.

Table 3. Evaluation Method used for Each Dataset.

Dataset	**RMSD**	**Cancer RMSD**	**Cancer node error**	**RMSE**	**Cancer RMSE**	**Dice Score**
Dataset 1-3	O	X	O	O	X	X
Dataset 4-7	O	O	X	O	O	O

3.1. Sensitivity Analysis with Dataset 4

The sensitivity analysis was conducted using Dataset 4, as it represents the most realistic assumptions among the breast phantom datasets. The RMSD of Dataset 4 is 1.11mm, and the cancer RMSD of Dataset 4 is 1.53mm. The basic setting is as followed: 8 layers of GraphSAGE [17] with max aggregation function, 80-channel, 100 random edge addition. Now for each experiment, we changed some properties from basic setting. The result of sensitivity analysis is analyzed in Table 4 and Fig 8, 9.

First, we examined the effect of varying the type of GNN model used. Specifically, we compared the performance of GCN [18], GraphSAGE, and GraphConv [19]. GCN is one of basic GNN model that relies on all edges for information propagation. On the other hand, GraphSAGE performs well in inductive learning scenarios by sampling neighbors. GraphConv is similar to GCN, but integrates neighbor information and self-information with separate weights. GraphSAGE shows the best performance in average cancer RMSE.

Next, we explored how varying the number of layers and channels affected the model's performance. When using too few layers, the distance required for information propagation can be insufficient. On the other hand, too many layers can cause oversmoothing, where node features become indistinguishable. The best performance was achieved with 8 layers.

For number of channels, a small number of channels reduced the model's capacity to capture complex features. Conversely, using too many channels can cause overfitting, again leading to increased average cancer RMSE. The best performance was achieved with 80 channels.

In addition, we examined the influence of random edge addition. These connections played a crucial role in alleviating the limitations of information propagation distance. However, when too many random edges are added, the structural information of the graph can be distorted. The best performance was achieved with 100 edges.

Finally, we tried to use jumping knowledge (JK) [20] for basic setting and 10-layer setting. JK is typically employed to mitigate oversmoothing in deep GNNs by aggregating node representations from multiple layers, allowing the model to retain information from earlier layers. The structure of JK used in our experiments is illustrated in Fig 10, and the pooling method applied was max pooling. However, contrary to the expected benefits, our results showed that using JK led to higher RMSE compared to the original configurations without JK.

Table 4. Result Tables of Sensitivity Analysis. The basic setting is as followed: 8-layer of GraphSAGE layer with max aggregation function, 80-channel, 100 random edge addition. Now for each experiment, we changed some properties from basic setting.

GNN	**RMSE** (mm)	**Cancer RMSE** (mm)	**Dice Score**
GCN	1.185 ± 0.454	1.586 ± 1.179	0.509 ± 0.129
GraphSAGE	0.932 ± 0.359	0.093 ± 0.056	0.925 ± 0.040
GraphConv	0.857 ± 0.397	0.099 ± 0.059	0.920 ± 0.041

Layers	**RMSE** (mm)	**Cancer RMSE** (mm)	**Dice Score**
6	0.876 ± 0.376	0.111 ± 0.062	0.913 ± 0.044
8	0.932 ± 0.359	0.093 ± 0.056	0.925 ± 0.040
10	1.045 ± 0.343	0.099 ± 0.057	0.922 ± 0.041
12	0.893 ± 0.343	0.115 ± 0.074	0.909 ± 0.050

Channels	**RMSE** (mm)	**Cancer RMSE** (mm)	**Dice Score**
64	0.922 ± 0.343	0.112 ± 0.060	0.911 ± 0.043
72	1.008 ± 0.333	0.112 ± 0.065	0.913 ± 0.045
80	0.932 ± 0.359	0.093 ± 0.056	0.925 ± 0.040
96	0.888 ± 0.391	0.100 ± 0.058	0.921 ± 0.041

Random Edge Addition	**RMSE** (mm)	**Cancer RMSE** (mm)	**Dice Score**
0	1.144 ± 0.380	0.831 ± 0.607	0.762 ± 0.089
100	0.932 ± 0.359	0.093 ± 0.056	0.925 ± 0.040
200	0.893 ± 0.375	0.104 ± 0.065	0.918 ± 0.046

Number of Layers with Add	**RMSE** (mm)	**Cancer RMSE** (mm)	**Dice Score**
0	0.932 ± 0.359	0.093 ± 0.056	0.925 ± 0.040
1	0.888 ± 0.387	0.089 ± 0.059	0.928 ± 0.043
2	0.857 ± 0.359	0.094 ± 0.056	0.924 ± 0.041

Layers	**RMSE** (mm)	**Cancer RMSE** (mm)	**Dice Score**
8	0.932 ± 0.359	0.093 ± 0.056	0.925 ± 0.040
8 w/ JK	0.876 ± 0.386	0.100 ± 0.062	0.920 ± 0.043
10	1.045 ± 0.343	0.099 ± 0.057	0.922 ± 0.041
10 w/ JK	0.928 ± 0.356	0.102 ± 0.065	0.918 ± 0.046

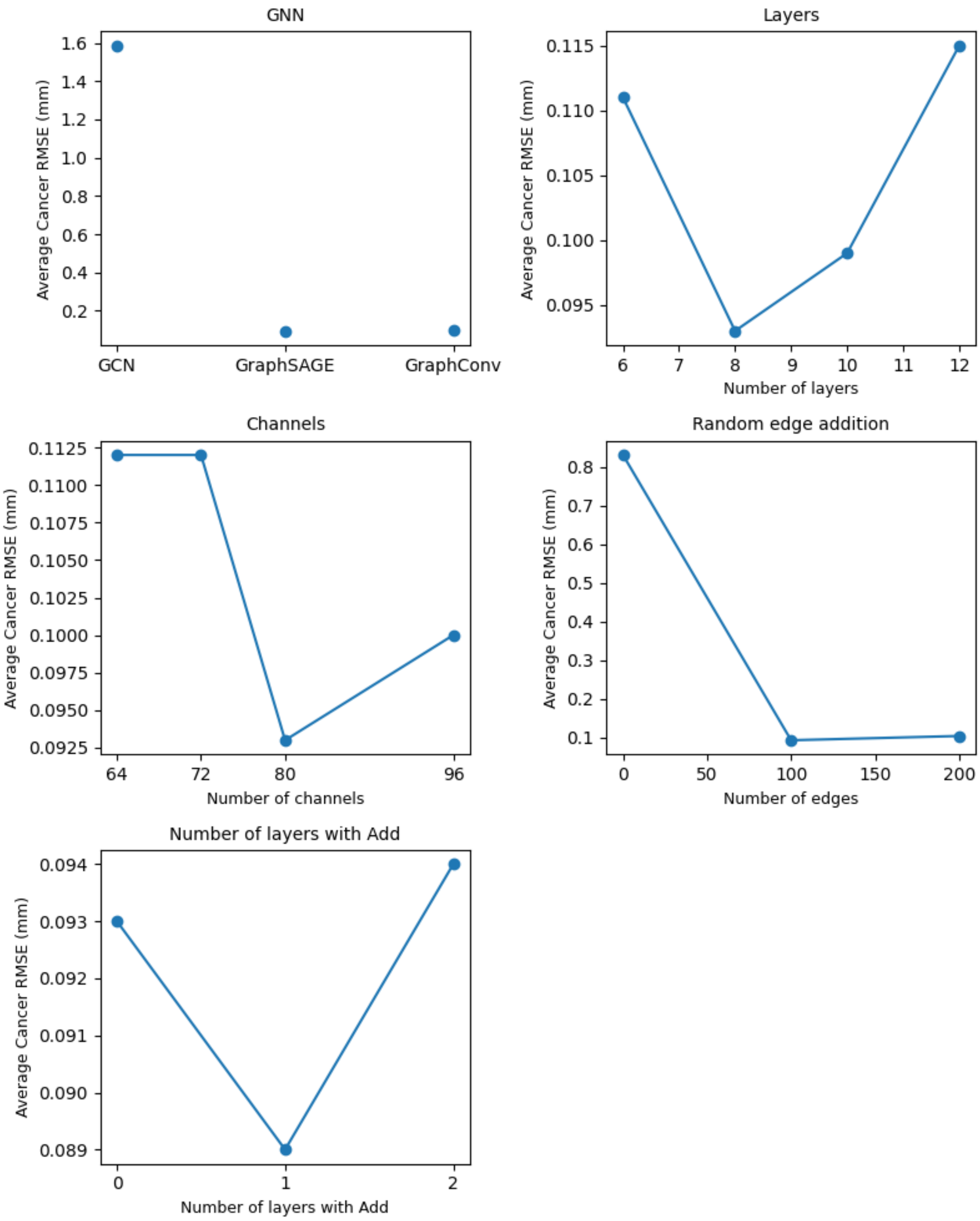


Fig 8. Cancer RMSE Graphs of Sensitivity analysis. The basic setting is as followed: 8-layer of GraphSAGE layer with max aggregation function, 80-channel, 100 random edge addition. Now for each experiment, we changed some properties from basic setting.

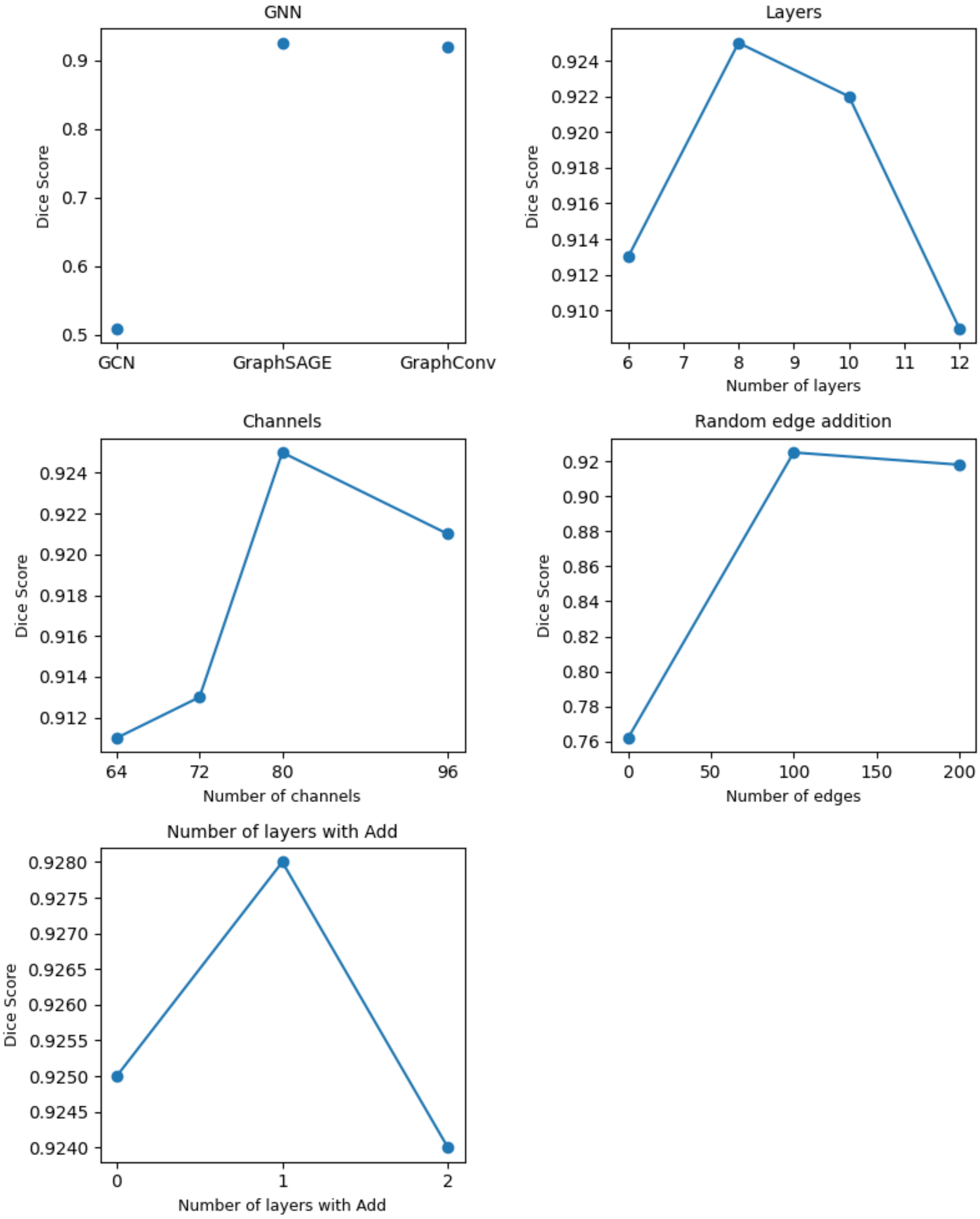


Fig 9. Dice Score Graphs of Sensitivity analysis. The basic setting is as followed: 8-layer of GraphSAGE layer with max aggregation function, 80-channel, 100 random edge addition. Now for each experiment, we changed some properties from basic setting.

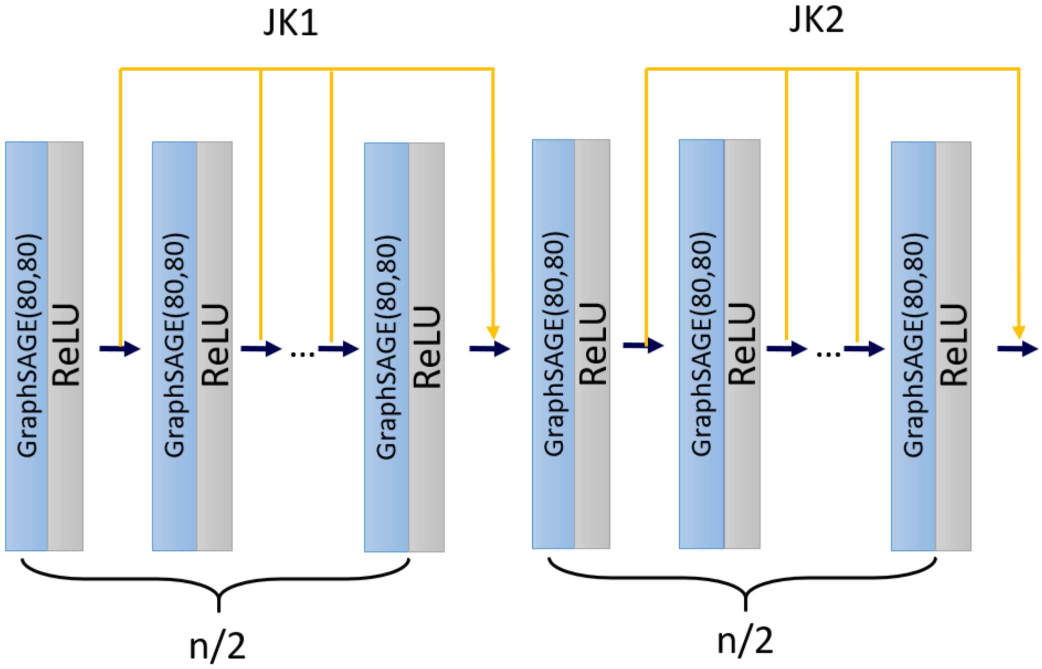


Fig 10. Jumping Knowledge Structure. The pooling method is max. If n=8, jumping knowledge process features of 4 layers and transfer to next layer. If n=10, jumping knowledge process features of 5 layers and transfer to next layer.

3.2. Performance of the GNN Model in Predicting Breast Cancer Site for Breast Phantom Dataset

Dataset 1

The specific model used is as follows: 6 layers of GraphSAGE, max aggregation, 80 channels, and 100 random edge additions. The detailed structure of this model is illustrated in Fig 11. This model was applied not only to Dataset 1 but also to Dataset 2 and 3. The performance metrics for this model include an RMSD of 17.41 mm, an average RMSE of 2.62 mm, and an average cancer node error of 1.78 mm.

Dataset 2

The same model as Dataset 1 was used, resulting in an RMSD of 11.78 mm and an average RMSE of 2.02 mm. The average cancer node error was 4.43 mm.

Dataset 3

The same model as Dataset 1 was used, resulting in an RMSD of 0.989 mm and an average RMSE of 0.173 mm. The average cancer node error was 0.350 mm.

Dataset 4

The specific model used is as follows: 8 layers of GraphSAGE, max aggregation except the first add aggregation layer, 80-channel, 100 random edge addition. The detailed structure of this model is shown in Fig 12. This model was also used for Dataset 5–7. For Dataset 4, the average RMSD was 1.11 mm, and the average cancer RMSD was 1.53 mm. The average RMSE was 0.888 mm, with an average cancer RMSE of 0.089 mm. The dice score for the cancer mass was 0.928.

Fig 13 illustrates the difference between ground truth and prediction for one example of test dataset. The black points are location of nodes before deformation. Colored points are location of nodes after deformation. The blue points are the ground truth simulated by ANSYS. The red points are the prediction produced by the model. All displacements have been magnified by 50 times for easier visual recognition.

Table 5 shows the result of Dataset 1-4.

Table 5. The Result of Dataset 1-4.

Dataset	**RMSD** (mm)	**RMSE** (mm)	**Cancer Error** (mm)
Dataset 1	17.41	2.62 ± 1.03	1.78 ± 0.95
Dataset 2	11.78	2.02 ± 1.01	4.43 ± 0.32
Dataset 3	0.989	0.173 ± 0.079	0.350 ± 0.218

Dataset	**RMSD** (mm)	**Cancer RMSD** (mm)	**RMSE** (mm)	**Cancer RMSE** (mm)	**Dice Score**
Dataset 4	1.11	1.53	0.888 ± 0.387	0.089 ± 0.059	0.928 ± 0.043

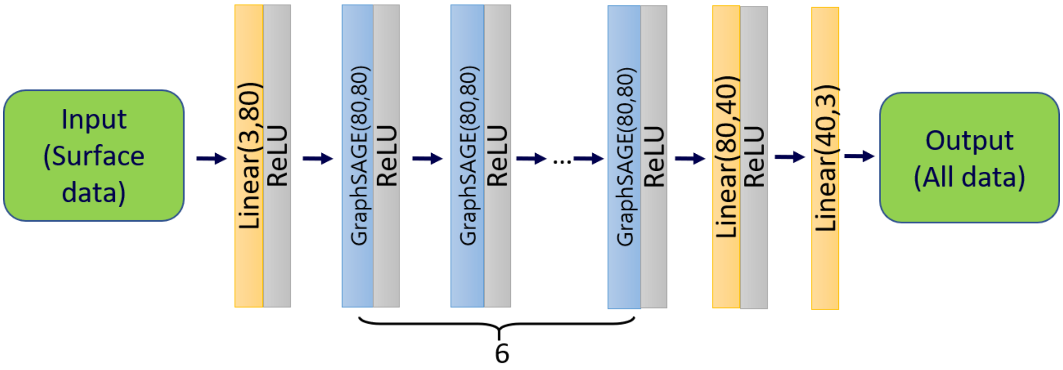


Fig 11. The Model used for Dataset 1-3. It consists of 6 of GraphSAGE layers and 3 linear layers. For aggregation function, max aggregation is used.

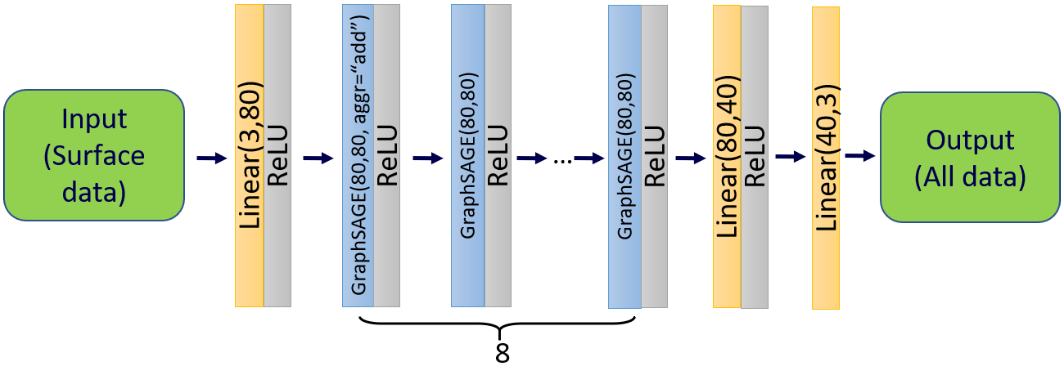


Fig 12. The Model used for Dataset 4-7. It consists of 8 of GraphSAGE layers and 3 linear layers. For aggregation function, max aggregation is used except the first add aggregation layer.

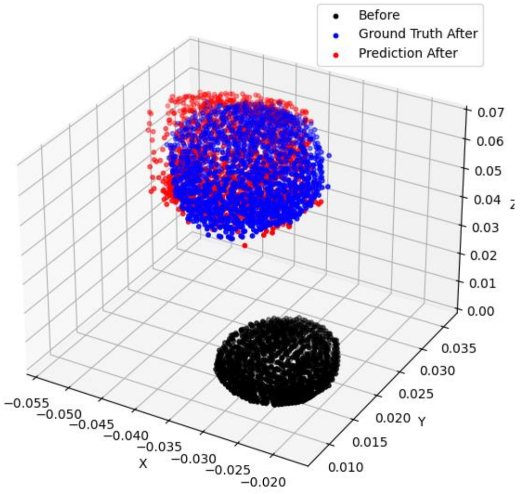


Fig 13. Cancer Sites of One Example in Dataset 4. These images illustrate the difference between ground truth and prediction. The black points are location of nodes before deformation. Colored points are location of nodes after deformation. The blue points are the ground truth simulated by ANSYS. The red points are the prediction produced by the model. All displacements have been magnified by 50 times for easier visual recognition.

3.3. Performance of the GNN Model in Predicting Breast Cancer Site for Real Patient Dataset

For Dataset 5-7, the same model is used as in Dataset 4.

Dataset 5

The average RMSD was 3.24 mm, and the average cancer RMSD was 3.52 mm. The average RMSE was 3.08 mm, with an average cancer RMSE of 0.31 mm. The dice score for the cancer mass was 0.683.

Dataset 6

The average RMSD was 1.50 mm, and the average cancer RMSD was 1.31 mm. The average RMSE was 7.71 mm, with an average cancer RMSE of 0.21 mm. The Dice score for the cancer mass was 0.827.

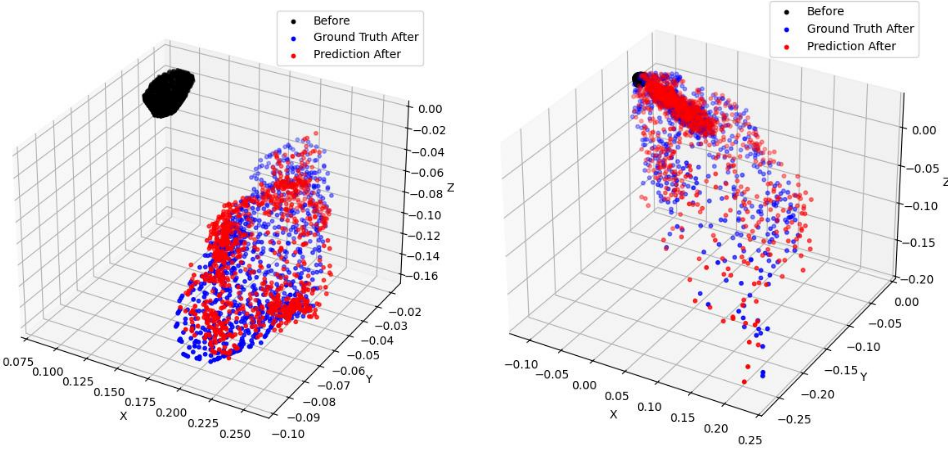
Dataset 7

The average RMSD was 1.23 mm, and the average cancer RMSD was 1.15 mm. The average RMSE was 1.10 mm, with an average cancer RMSE of 0.076 mm. The Dice score for the cancer mass was 0.911.

Table 6 shows the result of Dataset 5-7. In addition, Fig 14 illustrates the difference between ground truth and prediction for one example of test dataset in each dataset. The black mass is location of nodes before deformation. Colored masses are location of nodes after deformation. The blue mass is the ground truth simulated by ANSYS. The red mass is the prediction produced by the best model. All displacements have been magnified by 50 times for easier visual recognition.

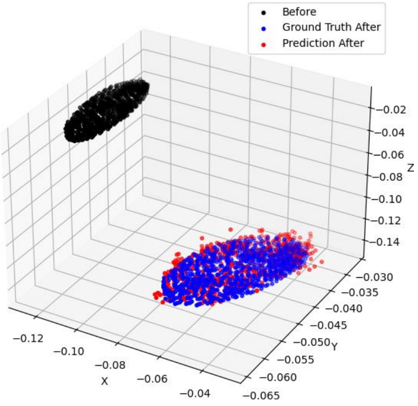
Table 6. The Result of Dataset 5-7.

Dataset	**RMSD (mm)**	**Cancer RMSD (mm)**	**RMSE (mm)**	**Cancer RMSE (mm)**	**Dice Score**
Dataset 5	3.24	3.52	3.08 ± 1.18	0.31 ± 0.18	0.683 ± 0.133
Dataset 6	1.50	1.31	7.71 ± 0.58	0.21 ± 0.11	0.827 ± 0.067
Dataset 7	1.23	1.15	1.11 ± 1.04	0.094 ± 0.108	0.891 ± 0.099



(a) Dataset 5

(b) Dataset 6



(c) Dataset 7

Fig 14. Cancer Sites of One Example in Test Dataset for Each Dataset. These images illustrate the difference between ground truth and prediction. The black points are location of nodes before deformation. Colored points are location of nodes after deformation. The blue points are the ground truth simulated by ANSYS. The red points are the prediction produced by the model. All displacements have been magnified by 50 times for easier visual recognition.

4. Discussion

4.1. Interpretation of Sensitivity analysis

In the first set of experiments, we evaluated different types of GNNs. GraphSAGE achieved the best performance, while GraphConv produced comparable results. However, GCN performed poorly. This discrepancy is likely due to GCN's approach of treating all edges equally during message passing, applying the same weight to self-information. In our graph structure, the importance of neighboring nodes varies significantly. Specifically, nodes closer to the surface tend to carry more critical information. GCN's use of all edges indiscriminately leads to a loss of accuracy, whereas GraphSAGE improves performance by selectively sampling edges. Moreover, in our dataset, both the input and output features represent node displacements, making self-information particularly important. GraphConv handles this effectively by using separate weights to combine neighbor information and self-information, a capability GCN lacks. Consequently, GCN suffers from reduced accuracy in this regard as well.

Next, we evaluated the model's performance across varying numbers of layers and channels, identifying 8 layers and 80 channels as the optimal configuration. Regarding the number of layers, models with fewer than 8 layers exhibited lower accuracy, likely due to limited information propagation across the graph. However, the decrease in accuracy when using more than 8 layers may not be attributed solely to oversmoothing. We will further explore this issue in the context of JK in a later paragraph. For the number of channels, models with fewer than 80 channels suffered from insufficient representational capacity, leading to lower accuracy—an observation consistent with general trends in neural networks. On the other hand, models with more than 80 channels also experienced a decline in accuracy, likely due to overfitting. A phenomenon that the increased complexity from excessive channels appears to hinder the model’s generalization ability is commonly observed across neural networks also.

In addition, we evaluated the influence of random edge addition. The best performance was achieved with 100 edges. Without any random edge addition, the result was bad. In this case, the information delivery distance is limited to a distance of up to $d \times n$ because only edges exist according to the edge condition we defined. (d : edge condition distance, n : number of layers) Therefore, message passing does not occur properly to the cancer nodes. With 200 edges, the model performed worse than with 100 edges. This is likely because random edge addition disregards the physical structure of the breast. While the existing edges indicate that connected nodes are spatially close, the randomly added edges lack this spatial relationship. Although these random edges may facilitate message passing to cancerous nodes, they inadvertently disrupt the structural integrity of the graph. As a result, adding too many random edges negatively impacts the model's performance.

In the final part of the sensitivity analysis, we evaluated the performance of Jumping Knowledge (JK). We expected JK to alleviate oversmoothing and yield better results. However, the actual results did not align with our expectations. Several factors may explain this outcome. First expected cause is that the JK implementation in our model was relatively simple. In PhysGNN [13], the authors used JK with LSTM pooling, which may have improved performance. However, due to the excessive training time required for LSTM, we opted not to use it. It's possible that LSTM pooling would have provided better results than the max pooling we used. Additionally, there are various types of JK, and the JK used in our model is just one example. There may be a more suitable JK structure that could better alleviate oversmoothing in our model. The second expected cause is that oversmoothing may not be the sole factor causing poor performance with many layers. If the JK mechanism in our model had been effective, it would have alleviated oversmoothing. However, the lack of improvement despite JK's potential suggests that other factors may be contributing to the poor results.

4.2. Interpretation of Results for Breast Phantom Dataset

Compared to RMSD, Dataset 2 and 3 showed relatively larger errors for cancer nodes than Dataset 1. This indicates that the model performs worse in scenarios involving varying force locations. However, in Dataset 4, cancer RMSE is not so big compared to cancer RMSD. This result may be interpreted as an impact of the model's complexity; however, there is another factor worth considering. While training on Dataset 4, we applied a train mask to the model. This may have significantly reduced the cancer RMSE, but coincidentally increased the overall RMSE. In fact, we observe that while the RMSE is poor in Dataset 4, the cancer RMSE is relatively good.

4.3. Analysis of Results for Real Patient Dataset

Our model is specifically tuned for Dataset 4, so we expected worse results for Dataset 5-7, and this was indeed the case. Several factors likely contributed to this. First, the volume of the breast and cancer mass differs from that in Dataset 4. Second, the location of the cancer mass is also different. These variations affect the graph structure, making it harder for our model to achieve optimal results.

Nevertheless, our model still performed reasonably well on Dataset 5-7. Despite the differences in graph structure, all datasets share a common property—they are based on the breast structure. This could explain why our model showed relatively good results for Dataset 5-7, even considering the above-mentioned factors.

4.4. Future Directions

In this study, we developed a GNN model that predicts the displacement of all nodes based on the displacement of surface nodes as input features. However, there may be other factors that could be used as input features. For example, incorporating the material properties of the nodes or their pre-deformation positions could be potential options. By including more information, the model's performance could likely improve, so exploring these possibilities could be beneficial.

Additionally, with the current approach, the model is trained on data from a single patient and must be applied to that specific patient, which introduces the drawback of requiring the patient to wait during the training process. Since GraphSAGE excels in inductive settings, it could be developed to learn from multiple patients' data and make predictions without the need for additional training. Alternatively, methods like LoRA (Low-rank adaptation) [21] could be explored to minimize the additional training required, thus reducing the wait time for patients.

5. Conclusion

In this study, a 3D finite element model of a breast was constructed using MRI data of breast phantom or real breast cancer patient. Using finite element simulation data from MRI, we generated datasets with varying material properties and force conditions. We created a dataset where the input consists of the displacements of the top surface nodes, and the output consists of the displacements of all nodes throughout the model. We then trained a GNN model using this dataset. GNN model used special graph structure that can represent physical structure. Finally, the GNN model succeeded to predict the site of breast cancer with high accuracy in real-time.

References

[1] F. Bray, J. Ferlay, I. Soerjomataram, R. L. Siegel, L. A. Torre, and A. Jemal, “Global cancer statistics 2018: Globocan estimates of incidence and mortality worldwide for 36 cancers in 185 countries,” CA: *a cancer journal for clinicians*, vol. 68, no. 6, pp. 394–424, 2018.

[2] R. A. Smith, D. Saslow, K. A. Sawyer, W. Burke, M. E. Costanza, W. P. Evans III, R. S. Foster Jr, E. Hendrick, H. J. Eyre, and S. Sener, “American cancer society guidelines for breast cancer screening: update 2003,” CA: *a cancer journal for clinicians*, vol. 53, no. 3, pp. 141–169, 2003.

[3] R. M. Mann, N. Cho, and L. Moy, “Breast mri: state of the art,” *Radiology*, vol. 292, no. 3, pp. 520–536, 2019.

[4] Mayo Clinic. 2023. Breast Biopsy. https://www.mayoclinic.org/tests-procedures/breast-biopsy/about/pac-20384812

[5] K. S. Myers, I. R. Kamel, and K. J. Macura, “Mri-guided breast biopsy: outcomes and effect on patient management,” *Clinical breast cancer*, vol. 15, no. 2, pp. 143–152, 2015.

[6] T. Alderliesten, C. Loo, A. Paape, S. Muller, E. Rutgers, M.-J. V. Peeters, and K. Gilhuijs, “On the feasibility of mri-guided navigation to demarcate breast cancer for breast-conserving surgery,” *Medical physics*, vol. 37, no. 6 Part1, pp. 2617–2626, 2010.

[7] M. Choi, M. Jang, S.-S. Yoo, G. Noh, and K. Yoon, “Deep neural network for navigation of a single-element transducer during transcranial focused ultrasound therapy: proof of concept,” *IEEE J. Biomed. Health Inform.*, vol. 26, no. 11, pp. 5653–5664, 2022.

[8] L. Liang, M. Liu, C. Martin, and W. Sun, “A deep learning approach to estimate stress distribution: a fast and accurate surrogate of finite-element analysis,” *J. R. Soc. Interface.*, vol. 15, no. 138, p. 20170844, 2018.

[9] N. Black and A. R. Najafi, “Learning finite element convergence with the multi-fidelity graph neural network,” *Computer Methods in Applied Mechanics and Engineering*, vol. 397, p.115120, 2022.

[10] M. V. Shivaditya, J. Alves, F. Bugiotti, and F. Magoul`es, “Graph neural network-based surrogate models for finite element analysis,” in *2022 21st International Symposium on Distributed Computing and Applications for Business Engineering and Science (DCABES)*, pp. 54–57, IEEE, 2022.

[11] X. Fu, F. Zhou, D. Peddireddy, Z. Kang, M. B.-G. Jun, and V. Aggarwal, “An finite element analysis surrogate model with boundary oriented graph embedding approach for rapid design,” *Journal of Computational Design and Engineering*, vol. 10, no. 3, pp. 1026–1046, 2023.

[12] C. Jiang and N.-Z. Chen, “Graph neural networks (gnns) based accelerated numerical simulation,” *Engineering Applications of Artificial Intelligence*, vol. 123, p. 106370, 2023.

[13] Y. Salehi and D. Giannacopoulos, “Physgnn: A physics-driven graph neural network based model for predicting soft tissue deformation in image-guided neurosurgery,” 2022.
[14] K.-J. Bathe, *Finite element procedures*. Klaus-Jurgen Bathe, 2006.

[15] H. Yin, L. Sun, G. Wang, T. Yamada, J. Wang, and M. Vannier, “Imageparser: a tool for finite element generation from three-dimensional medical images,” *BioMedical Engineering OnLine*, vol. 3, pp. 1–9, 2004.

[16] J. Zhou, G. Cui, S. Hu, Z. Zhang, C. Yang, Z. Liu, L. Wang, C. Li, and M. Sun, “Graph neural networks: A review of methods and applications,” *AI open*, vol. 1, pp. 57–81, 2020.

[17] W. Hamilton, Z. Ying, and J. Leskovec, “Inductive representation learning on large graphs,” *Advances in neural information processing systems*, vol. 30, 2017.

[18] T. N. Kipf and M. Welling, “Semi-supervised classification with graph convolutional networks,” *arXiv preprint arXiv:1609.02907*, 2016.

[19] C. Morris, M. Ritzert, M. Fey, W. L. Hamilton, J. E. Lenssen, G. Rattan, and M. Grohe, “Weisfeiler and leman go neural: Higher-order graph neural networks,” in *Proceedings of the AAAI conference on artificial intelligence*, vol. 33, pp. 4602–4609, 2019.

[20] K. Xu, C. Li, Y. Tian, T. Sonobe, K.-i. Kawarabayashi, and S. Jegelka, “Representation learning on graphs with jumping knowledge networks,” in *International conference on machine learning*, pp. 5453–5462, PMLR, 2018.

[21] E. J. Hu, Y. Shen, P. Wallis, Z. Allen-Zhu, Y. Li, S. Wang, L. Wang, and W. Chen, “Lora: Low-rank adaptation of large language models,” 2021.

Abstract in Korean

그래프 신경망을 이용한 실시간 유방암 위치 예측

유방암은 여성의 암 관련 사망 원인 중 제일 큰 비중을 차지하지만, 조기 진단이 치료 결과를 상당히 개선할 수 있다. 유방 MRI 는 유방암을 조기에 발견하는데 중요한 역할을 하며, 특히 X-ray 와 초음파에서 발견되지 않는 병변을 찾아내는 데 유용하다. 그러나 유방암의 확진을 위해서는 biopsy(생검)이 필요하며, 이 중 Core needle biopsy 가 가장 일반적인 방법이다. 그 중 direct MRI-guided breast biopsy 는 매우 정확하지만 시간이 많이 소요되고, 비용이 많이 든다는 단점이 있다. 이러한 한계를 해결하기 위해 최근 indirect MRI-guided breast biopsy 시스템이 제안되었으며, 이는 MRI 유도 내비게이션을 활용하여 MRI 방 밖에서 biopsy 를 수행하는 방식을 뜻한다. 그러나 기존 내비게이션 시스템은 실시간으로 조직 변형을 고려할 수 없는 한계가 있다. 이 문제를 해결하기 위해 그래프 신경망(Graph Neural Network, GNN)을 사용하여 실시간 3D 변형 모델을 개발했다. 이 모델은 표면 변위 정보를 토대로 유방 조직 변형을 예측한다. 모델 학습을 위해 유방 팬텀 및 실제 환자 MRI 데이터로 생성된 3D 모델에 대한 유한 요소 시뮬레이션을 사용하여 다양한 물질 특성과 힘 조건을 가진 데이터셋을 만들었다. 이러한 데이터셋을 기반으로, 물리적 구조를 표현할 수 있는 특별한 그래프 구조를 사용하여 모든 노드의 변위를 표면 노드의 변위에서 예측하도록 모델을 학습시켰다. 그 결과 최적의 모델을 사용하여 평균 RMSE 0.888 mm 및 평균 암 RMSE 0.089 mm, 그리고 dice score 0.928 을 달성했다. 또한, 동일한 모델을 사용하여 실제 환자 데이터를 학습시킨 결과 좋은 성능을 보였다.

핵심되는 말 : GNN, Biopsy(생검), 유방암, MRI