IOP Publishing Phys. Med. Biol. 65 (2020) 065011 https://doi.org/10.1088/1361-6560/ab7505

Physics in Medicine & Biology

IPEM Institute of Physics and Engineering in Medicine

PAPER

RECEIVED
13 November 2019

REVISED
27 January 2020

ACCEPTED FOR PUBLICATION
11 February 2020

PUBLISHED
20 March 2020

Physics-guided machine learning for 3-D quantitative quasi-static elasticity imaging

Cameron Hoerig[1,2], **Jamshid Ghaboussi**[3] **and Michael F Insana**[1,2]

[1] Department of Bioengineering, University of Illinois at Urbana-Champaign, Urbana, IL 61801 United States of America
[2] Beckman Institute of Advanced Science and Technology, University of Illinois at Urbana-Champaign, Urbana, IL 61801 United States of America
[3] Department of Civil and Environmental Engineering, University of Illinois at Urbana-Champaign, Urbana, IL 61801 United States of America

E-mail: hoerig2@illinois.edu **and** choerig@riversideresearch.org.

Keywords: machine learning, elastography, 3D imaging, inverse problems

Abstract

We present a 3D extension of the Autoprogressive Method (AutoP) for quantitative quasi-static ultrasonic elastography (QUSE) based on sparse sampling of force-displacement measurements. Compared to current model-based inverse methods, our approach requires neither geometric nor constitutive model assumptions. We build upon our previous report for 2D QUSE and demonstrate the feasibility of recovering the 3D linear-elastic material property distribution of gelatin phantoms under compressive loads. Measurements of boundary geometry, applied surface forces, and axial displacements enter into AutoP where a Cartesian neural network constitutive model (CaNNCM) interacts with finite element analyses to learn physically consistent material properties with no prior constitutive model assumption. We introduce a new regularization term uniquely suited to AutoP that improves the ability of CaNNCMs to extract information about spatial stress distributions from measurement data. Results of our study demonstrate that acquiring multiple sets of force-displacement measurements by moving the US probe to different locations on the phantom surface not only provides AutoP with the necessary information for a CaNNCM to learn the 3D material property distribution, but may significantly improve the accuracy of the Young's modulus estimates. Furthermore, we investigate the trade-offs of decreasing the contact area between the US transducer and phantom surface in an effort to increase sensitivity to surface force variations without additional instrumentation. Each of these modifications improves the ability of CaNNCMs trained in AutoP to learn the spatial distribution of Young's modulus from force-displacement measurements.

1. Introduction

Quantitative elasticity imaging of breast tissue *in vivo* is inherently a three-dimensional (3D) inverse problem. Ophir *et al* first described quasi-static ultrasonic elastography (QUSE) whereby radio frequency (RF) echo frames are acquired as an ultrasound (US) transducer is slowly pressed into the surface of tissue (Ophir *et al* 1991). Application of a correlation-based speckle-tracking algorithm to pre- and post-deformation frames provides estimates of tissue displacements within the field of view. Calculating gradients of the displacements along the beam axis to estimate strain fields is a qualitative form of QUSE often referred to as 'strain imaging'. Quantitative estimates of tissue mechanical properties require the displacement or strain measurements be input to an inverse problem to estimate parameters of a pre-selected constitutive model. In these model-based approaches, stresses can be eliminated using the selected constitutive model or assumed to be approximately uniform when a large compressor applies the force load (Ophir *et al* 1991).

Clinical US breast imaging exams are typically performed with one-dimensional (1D) linear arrays, meaning the B-mode images—and displacement estimates—are two-dimensional (2D) (Szabo and

© 2020 Institute of Physics and Engineering in Medicine

Lewin 2013). Furthermore, displacements can be estimated sensitively only in the direction of wave propagation where signal phase is available, thus limiting measurements to 1-D (axial) displacements in a 2D plane (Lubinski *et al* 1999). Even though 2D QUSE has demonstrated feasibility for improving detection and differential diagnosis of malignant breast lesions (Schaefer *et al* 2011, Zhi *et al* 2007, Athanasiou *et al* 2010, Liu *et al* 2016, Barr 2010, Barr *et al* 2012), improper but often necessary 2D plane-stress or plane-strain assumptions in the inverse problem may lead to artifacts when the assumptions are violated, thus limiting the diagnostic value of elastograms (Doyley 2012).

Several methods for estimating strain or material parameters within a 3D tissue volume have been proposed. Perhaps the most straight-forward technique is a free-hand sweep of a 1-D linear array in the elevational direction to estimate displacements at regularly spaced intervals throughout the volume. The set of RF echo frames can be 'stitched' together into a 3D data set upon which speckle-tracking algorithms can operate to estimate 3D displacements all through the scanned volume (Foroughi *et al* 2013). Alternatively, speckle-tracking methods can operate on pairs of pre- and post-compression frames at each location. Axial strains computed as the gradient of the each displacement estimate can be concatenated to form a 3D strain volume (Lindop *et al* 2006, Lee *et al* 2018). The latter fails to make use of 3D echo data whereas the former suffers from severe decorrelation effects arising from uncertainty in estimated probe position, a common problem in swept synthetic aperture imaging (Bottenus *et al* 2016). Richards *et al* automated the data acquisition process by attaching a US transducer to a positioning system and collecting volumetric RF data after compression with platens (Richards *et al* 2009), greatly reducing jitter errors between successive RF frames while utilizing full 3D echo information for displacement estimation. Similarly, other investigators adapted an automated breast volume scanning (ABVS) system to collect pre- and post-compression RF data (Wang *et al* 2017, Hendriks *et al* 2016, Hendriks *et al* 2018, Tyagi *et al* 2017). Even though these automated methods reduce decorrelation caused by uncertainty of the US probe position, they require long acquisition times and are thus susceptible to motion artifacts in displacements estimated via speckle-tracking. Other methods of acquiring 3D RF echo data include the use of a mechanically swept 1D array (Sayed *et al* 2013, Sayed *et al* 2014, Fisher *et al* 2007, Bharat *et al* 2008, Treece *et al* 2008, Housden *et al* 2011) or a 2D matrix array (Deprez *et al* 2009, Papadacci *et al* 2016b, Papadacci *et al* 2016a, Gijsbertse *et al* 2016, Fisher *et al* 2010). Compared to data acquisition via free-hand scanning or an ABVS system, these methods are less susceptible to motion artifacts between pre- and post-deformation frames. Generally, measurements for 3D QUSE are densely sampled to estimate material parameters throughout a continuous volume.

We previously developed 2D Cartesian neural network constitutive models (CaNNCMs) capable of learning the spatially varying stress-strain relationship of linear-elastic materials from force-displacement measurements without prior model assumptions (Hoerig *et al* 2019). Unlike other quantitative QUSE approaches, CaNNCMs are capable of learning material properties from a sparse sampling of measurement data. Considering that CaNNCMs provide a non-parametric description of mechanical behavior, further development of this method may provide the tools needed to discover which material properties are most relevant to clinical imaging simply by improving the accuracy of material property estimates. Accurate estimates provide device-independent results that can minimize instrumentation contributions to inter-patient variability.

Barbone and Gokhale previously proved the Young's modulus distribution can be estimated only up to a multiplicative constant using displacement measurements when force measurements are unavailable (Barbone and Gokhale 2004). Tyagi *et al* later proposed methods for incorporating force measurements into QUSE (Tyagi *et al* 2014). The force boundary conditions (BCs) considered in Tyagi *et al* (2014) were equivalent to a single axial force measured as a probe was pressed into part of the object surface; i.e. there were no measurements of the force distribution across the phantom surface. Similarly, in this report, we collect force-displacement measurements in independent planes throughout the volume of a heterogeneous gelatin cube as an ultrasound probe compressively loads one surface. As the load is applied, surface force is measured while pre- and post-deformation RF echo data are acquired. Measurements resulting from one compressive load step are the applied force and an estimate of axial displacements within the US image plane. The important distinction is that full interior displacement data were used in Tyagi *et al* (2014) whereas we only consider displacements within a plane for each compressive load. We will demonstrate how a single measurement of surface force and interior displacements limited to one plane can be a detriment to AutoP, but a new regularization term uniquely suited to AutoP can extract additional information encoded within measurement data to improve the ability of CaNNCMs to learn material properties. We then compare the effect of compressor size on Young's modulus estimates. Finally, we evaluate the effect of US probe force-displacement measurement errors on Young's modulus estimates and propose sampling strategies for effective 3D elastography using CaNNCMs trained in AutoP.

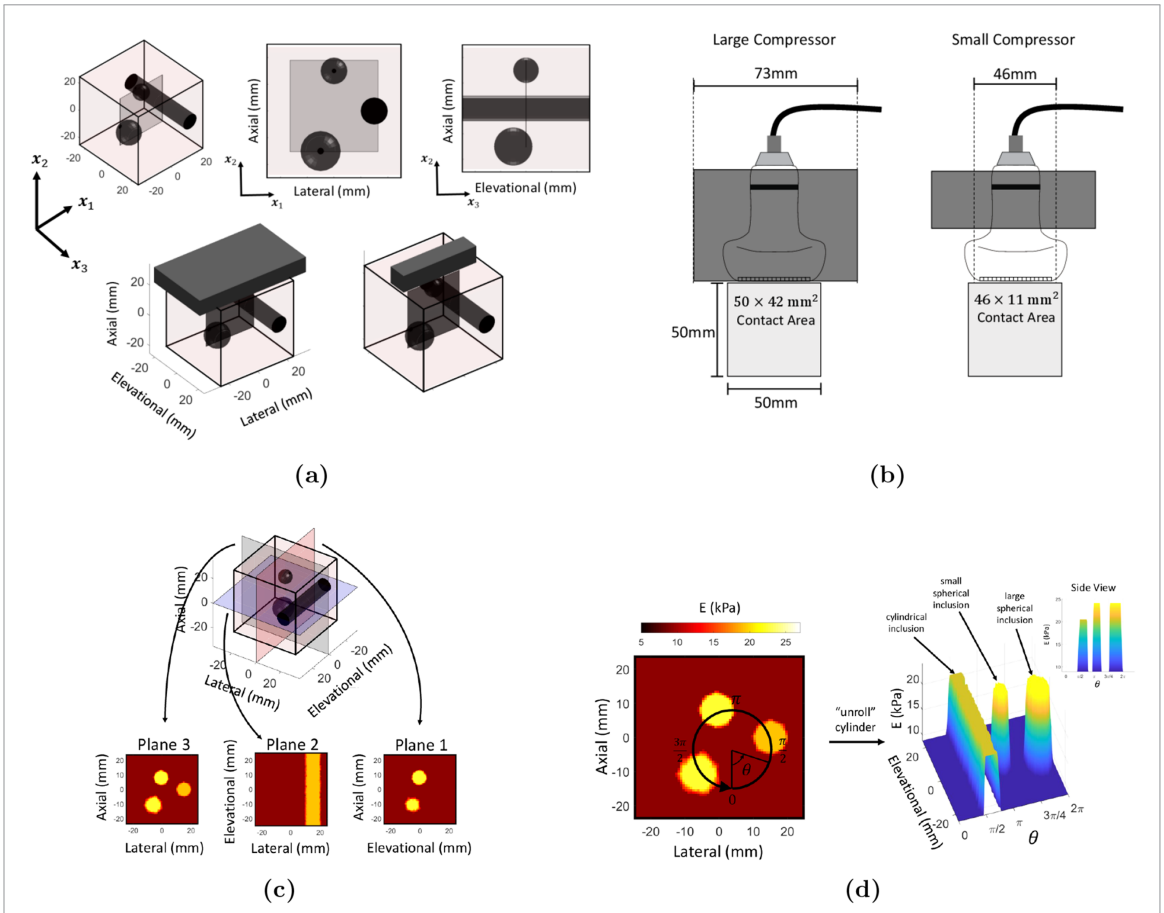

Figure 1. (a) Geometry of the linear-elastic gelatin phantom. Three stiff inclusions are embedded in a soft background material (8.9 ± 0.2 kPa). The spherical inclusions are the same stiffness (24.2 ± 0.8 kPa) whereas the cylindrical inclusion is slightly softer (20.5 ± 0.3 kPa). The bottom row compares the large and small probe compressor sizes relative to the phantom. Gray squares within the phantom indicate the scan plane of the US probe when centered in both the lateral and elevational directions relative to the beam axis. (b) The two compressor geometries are shown. (c) Representation of the 3D Young's modulus distribution within three orthogonal planes. Plane 3 (black) spans $x_1 - x_2$, Plane 2 (blue) spans $x_1 - x_3$, and Plane 1 (red) spans $x_2 - x_3$. The number for each plane was chosen to indicate the direction of the normal vector to the plane. For example, the normal vector to Plane 3 is $\boldsymbol{x}_3$. (d) An alternative method of visualizing the Young's modulus to observe the shape of the inclusions along all three spatial dimensions simultaneously. The circle circumscribed on Plane 3 (left) indicates the boundary of a cylinder that extends through the phantom in the x_3 direction. Computing the Young's modulus on this boundary and 'unrolling' the cylinder yields the surface plot on the right.

2. Experimental methods

2.1. Measurements on linear-elastic gelatin phantom

We manufactured a phantom using a mixture of deionized water, gelatin powder, and corn starch added to provide acoustic scattering. The phantom was a $50 \times 50 \times 50$ mm^3 cube comprised of a soft background material and three embedded inclusions. Two of the inclusions were spheres with radii of 5 and 7.5 mm whereas the third inclusion was a cylinder with 5 mm radius and 50 mm length. Separate macro-indentation measurements were performed on a sample of each material manufactured under the same conditions to provide an independent Young's modulus estimate (Altahhan *et al* 2016). For these indentation measurements, a 5 mm spherical probe was pressed into the sample surface up to 1 mm while acquiring force measurements. The force-displacement loading curves were fit to a Hertzian contact model to estimate the Young's modulus of the tissue sample. Indentation measurements were performed using a TA.XTPlus texture analyzer (Texture Technologies, Hamilton, MA). Using this procedure, the Young's modulus values of the background gelatin material, spherical inclusions, and cylindrical inclusion were estimated to be 8.9 ± 0.2 kPa, 24.2 ± 0.8 kPa, and 20.5 ± 0.3 kPa, respectively. A diagram of the phantom geometry is shown in figure 1(a). Images in the top row of figure 1(a) are three different views of the phantom.

Measurements were acquired using the same experimental setup described in (Hoerig *et al* 2017, Hoerig *et al* 2019). Compressive loads were applied by the US probe to induce 3% total strain over four load increments. Axial force, probe position and orientation, and a RF echo frame were acquired after each load increment. A Siemens Sonoline Antares ultrasound system (Siemens Healthcare USA, Mountainview, CA)

Table 1. Description of the force-displacement data sets. The field 'x_3 Position' indicates the offset of the US probe in elevation relative to the center plane of the phantom. The addition of 'S' before the set number indicates simulated data. Simulated data were limited to 3% strain, hence no maximum force values are reported for S1–S3, S9–S11 for 7% applied strain.

Set #	Phantom Rotation	x_3 Position	Max. Force (mN) (3% Strain)	Max. Force (mN) (7% Strain)	
1	0°	0 mm	943	2345	Large Compressor
2	0°	–4 mm	929	2337	
3	0°	+4 mm	1053	2545	
4	0°	–8 mm	881	2203	
5	0°	+8 mm	863	2230	
6	0°	–12 mm	825	2084	
7	0°	+12 mm	836	2141	
8	90°	0 mm	1012	2478	
9	0°	0 mm	509	1281	Small Compressor
10	0°	–4 mm	501	1284	
11	0°	+4 mm	481	1248	
12	0°	–8 mm	463	1214	
13	0°	+8 mm	455	1200	
14	0°	–12 mm	428	1131	
15	0°	+12 mm	417	1103	
16	90°	0 mm	518	1332	
S1	0°	0 mm	714	N/A	Simulated Large Compressor
S2	0°	–4 mm	697	N/A	
S3	0°	+4 mm	697	N/A	
S9	0°	0 mm	427	N/A	Simulated Small Compressor
S10	0°	–4 mm	419	N/A	
S11	0°	+4 mm	418	N/A	

with a VF10-5 linear array probe at 8 MHz center frequency was used for RF data collection. The speckle-tracking algorithm GLobal Ultrasound Elastography (GLUE) (Hashemi and Rivaz 2017) operated on the echo data to estimate axial displacements within the echo acquisition region. We refer to a data set as the collection of axial forces applied by the US probe, axial probe displacement, and internal displacements measured over all load increments in an experiment.

Two different rigid holders were used to attach the US probe to the positioning system. We denote these as large and small compressors (figure 1(b)). For the large compressor, the holder completely encased the probe and extended the transducer face in both x_1 and x_3 dimensions to 73×42 mm^2. Conversely, the small compressor probe holder encased only the transducer handle, leaving the face of the probe at its native 46×11 mm^2. Images in the bottom row of figure 1(a) compare the relative size of the large and small probe compressors with the phantom surface. Measurements acquired with both compressor sizes allowed us to explore how the shape of the applied force influences training and elasticity imaging accuracy.

Eight data sets were acquired with each compressor and are compiled in table 1. Before each set of measurements a small compressive pre-load was applied by the US transducer. Gravity loading and body forces caused a slight initial deformation of the phantom. The pre-load was intended to offset the effects of these body forces to provide a known initial configuration and ensure complete contact between the compressor and phantom surface. After the applying the pre-load, the phantom height was reduced to 48 mm. Data sets 1–7 were acquired by positioning the large compressor at $x_3 = 0$, –4, +4, –8, +8 mm, –12, and +12 mm. Each of these data sets was separated by at least 5 elevational beam widths, ensuring no correlation in the RF echo signals. Measurements for data set 8 were collected by rotating the phantom 90° about the x_3-axis and positioning the probe at $x_3 = 0$. Rotating the phantom in this way places the cylindrical inclusion closest to the probe. Data sets 9–16 were acquired in the same manner as sets 1–8 except the large compressor was exchanged for the small compressor.

Experiments 1–16 were repeated by increasing the magnitude of the applied load to induce 7% global strain over a total of six load increments. The extra two load increments were added to avoid decorrelation errors during speckle tracking and numerical instability in later finite element analyses (FEAs). Gelatin phantoms remain linear in their elastic response up to ≈10% strain (Hall *et al* 1997) and so there should be no distortion arising from material non-linearity. We denote these additional data sets using the same numbers in table 1 with an additional 'L'; e.g. data set 1 was acquired under 3% total strain whereas data sets 1 L implies 7% total strain.

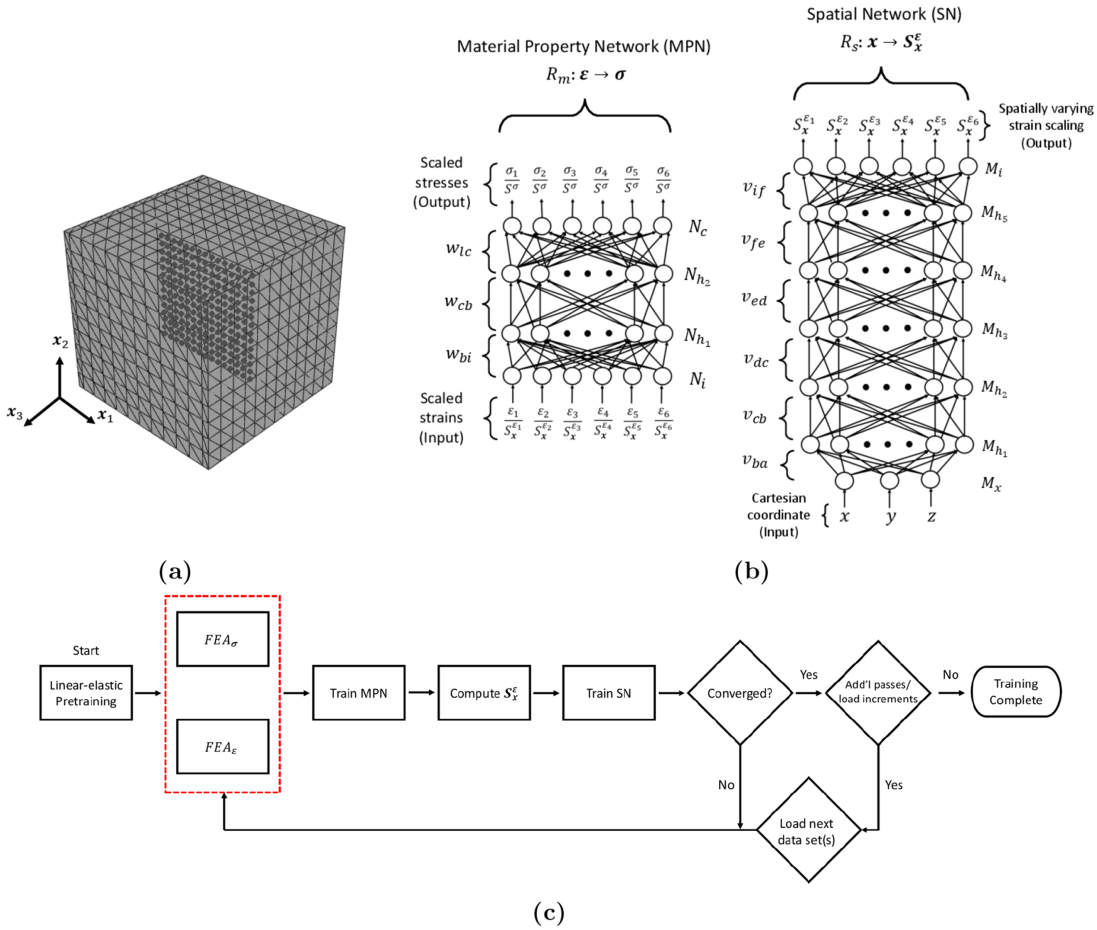

Figure 2. (a) Finite element mesh of the gelatin phantom used in AutoP and to generate simulated data sets. Interior nodes residing in an ultrasonic scan plane at $x_3 = 0$ are highlighted. (b) Architecture of a 3D CaNNCM. Variables v and w indicate network weights, while N and M are numbers of nodes in corresponding layers. Stresses and strains are expressed in Voigt notation and map from the original tensors as $[\sigma_{11}\ \sigma_{22}\ \sigma_{33}\ \sigma_{12}\ \sigma_{13}\ \sigma_{23}] \rightarrow [\sigma_1\ \sigma_2\ \sigma_3\ \sigma_4\ \sigma_5\ \sigma_6]$ and $[\varepsilon_{11}\ \varepsilon_{22}\ \varepsilon_{33}\ 2\varepsilon_{12}\ 2\varepsilon_{13}\ 2\varepsilon_{23}] \rightarrow [\varepsilon_1\ \varepsilon_2\ \varepsilon_3\ \varepsilon_4\ \varepsilon_5\ \varepsilon_6]$. The material property network (MPN) accepts a scaled strain vector as input and returns a scaled stress vector at the output. S_x^ε and S^σ are scale factors weighting the inputs and outputs of the MPN. Together, the MPN and spatial network (SN) learn the spatial distribution of material properties. (c) Flow diagram summarizing the AutoP training process briefly described in section 2.5. More details can be found in Hoerig *et al* (2019).

2.2. Finite element model

A finite element (FE) model of the gelatin phantom and US probe was created in ABAQUS 6.13 commercial finite element software to generate simulated, noise-free force-displacement data sets using a linear-elastic constitutive model with known Young's modulus and Poisson's ratio. This same mesh was later used to train CaNNCMs in AutoP using both simulated and experimentally acquired measurement data. The FE mesh used for experimental modeling is shown in figure 2(a). This mesh was comprised of 14043 10-node tetrahedral elements (20354 nodes). Points highlighted in figure 2(a) correspond to mesh nodes located in the scan plane when the US probe is positioned at $x_3 = 0$ mm. Internal displacement measurements applied in FEA_ε (described in section 2.5) are imposed at these nodes. We chose to model the phantom in its pre-strained state; thus, the phantom mesh spanned 51 mm in both the x_1 and x_3 directions but only 48 mm in x_2. The extension of the phantom in the lateral and elevational directions was to account for incompressibility of the gelatin phantom. Both the large and small compressor FE models consisted of eight 8-node hexahedral elements characterized by a very large shear modulus to approximate a rigid body. Contact between the probe and phantom FE meshes was defined to be frictionless and small sliding, the latter to increase stability of the contact mechanics. The bottom nodes of the phantom were constrained by a 'pinned' boundary condition for all FEAs. Any force BCs were applied as point loads to the top-middle node of the probe model. Motion of the probe was constrained to only the x_2 direction to avoid rigid body motion in the lateral or elevational directions.

In section 2.4, we will describe AutoP, wherein force and displacement BCs are applied in separate FEAs. Internal deformation of the phantom volume as estimated via speckle-tracking is limited to axial displacements within the scan plane. We measured displacements at up to seven parallel planes in the

phantom during AutoP training. To accommodate as many as seven ultrasound scan planes, we made sure there were meshed planes at $x_3 = -12, -8, -4, 0, +4, +8, +12$ mm. All FEAs were computed in ABAQUS 6.13.

Most FEAs in this study adopted the right-hand coordinate system shown in figure 1(a) where the origin is placed at the center of the phantom. All force-displacement measurements are acquired with reference to this fixed coordinate system (e.g. the axial direction is always along x_2 regardless of phantom rotation). However, the phantom was rotated before acquiring Data Sets 8 and 16. We account for this by rotating the phantom mesh to match experimental conditions before solving the relevant FEA. Then, computed stress-strain fields are rotated back to the fixed coordinate system before further processing.

2.3. Simulated noise-free data

Noise-free force displacement data sets were generated by simulating some of the experiments described in section 2.1. Using the Young's modulus values of the phantom estimated via macro-indentation, Poisson's ratio $\nu = 0.3$, and the previously described FE mesh[1], we first performed displacement controlled FEAs by applying displacement BCs to the US probe in the FE model to induce 3% compressive strain over four equal load increments (1.43 mm total axial displacement). We chose $\nu = 0.3$ to match the Poisson's ratio selected for CaNNCM pre-training, described in appendix A. Selecting a Poisson's ratio near or equal to 0.5 can cause difficulty in generating physically realizable stress-strain pre-training data and numerical instabilities in a FEA. As later shown in section 3, pre-training a CaNNCM as a compressible material does not inhibit the ability of network to accurately estimate Young's modulus values even for incompressible gelatin phantoms.

Contact forces were estimated in the displacement-controlled FEAs to simulate noise-free force values. Then, to ensure consistency in the simulated data, the computed contact forces were applied as a point load to the probe in separate FEAs to simulate noise-free displacements. This method was used to generate simulated versions of data sets 1–3 and 9–11. We label the simulated data sets S1–S3 and S9–S11 to remain consistent with the naming convention of the experimental data sets. Two additional data sets derived from S1 and S9 utilized axial displacements at all nodes in the phantom mesh. We label these S1′ and S9′ and their purpose is to test AutoP and 3D CaNNCMs under nearly ideal conditions.

2.4. Autoprogressive method with 3D CaNNCMs

The Autoprogressive (AutoP) method is a supervised machine learning technique that learns spatial distributions of stresses and strains as force-displacement measurements interact with finite-element algorithms and neural network structures (Ghaboussi *et al* 1998). AutoP is a fundamentally different approach to the inverse problem in elastography. Force-displacement measurements and knowledge of external object shape are combined with artificial neural networks (ANNs) and finite element analysis to iteratively produce increasingly accurate estimates of all stresses and strains free from prior constitutive model assumptions. The power of AutoP lies in the combination of physical modeling through FEA and machine learning that extracts stress-strain information from force-displacement measurements. By tightly coupling the ANNs with FEA, principles of mechanics—stress equilibrium and deformation compatibility conditions—constrain the material behavior learned by the networks to lead to physically consistent soft-computational constitutive models. After an ANN learns the stress-strain relationships, material parameters can be estimated from any chosen constitutive model applied in a separate process. Even though prior research has shown neural networks trained in AutoP can learn the mechanical behavior of fairly complex materials (e.g. Hashash *et al* (2006a), Hashash *et al* (2006b), Jung and Ghaboussi (2006), Jung and Ghaboussi (2010), Kim *et al* (2012)), tests of AutoP using CaNNCMs have so far been limited to linear-elastic materials.

As we described for 2D elasticity imaging (Hoerig *et al* 2019), CaNNCMs are comprised of a material property network (MPN) that learns a spatially averaged mapping from strain to stress ($R_m : \boldsymbol{\varepsilon} \rightarrow \boldsymbol{\sigma}$) and a spatial network (SN) that maps coordinate input to a spatially varying strain-scaling vector ($R_s : \boldsymbol{x} \rightarrow \boldsymbol{S}^{\varepsilon}_{\boldsymbol{x}}$). These two networks cooperate to learn the spatially-varying stress-strain relationship exhibited by an object. Figure 2(b) illustrates the 3D CaNNCM architecture.

[1] The FE mesh does not match the internal phantom geometry. Therefore, to create the spherical inclusions, we utilized a user-defined field (USDFLD) in Abaqus.

The MPN in a 3D CaNNCM has $N_c = 6$ components at the input and output. The lth component of a stress vector σ_l^n computed by the MPN for the nth input strain vector $\boldsymbol{\varepsilon}(\boldsymbol{x}_j)$ at spatial location $\boldsymbol{x}_j$ can be expressed as

$$\sigma_l^n(\boldsymbol{\varepsilon}(\boldsymbol{x}_j)) = S^{\sigma} \tanh\left(\sum_{c=1}^{N_{h_2}} w_{lc} \tanh\left(\sum_{b=1}^{N_{h_1}} w_{cb} \tanh\left(\sum_{i=1}^{N_i} w_{ba} \tanh\left(\frac{\varepsilon_i(\boldsymbol{x}_j)}{S_{\boldsymbol{x}_j}^{\varepsilon_i}}\right)\right)\right)\right), \tag{1}$$

where $1 \leq (l, i) \leq N_c = N_i = 6$. $N_{h_1} = N_{h_2} = 10$ are the number of nodes in the first and second hidden layers, respectively, and $\tanh(\cdot)$ denotes a hyperbolic tangent function. Spatial scaling values $S_{\boldsymbol{x}_j}^{\varepsilon_i}$ for the ith strain component are computed by the SN with input $\boldsymbol{x}_j$,

$$S_{\boldsymbol{x}_j}^{\varepsilon_i} = \phi\left(\sum_{f=1}^{M_{h_5}} v_{if} \tanh\left(\sum_{e=1}^{M_{h_4}} v_{fe} \tanh\left(\sum_{d=1}^{M_{h_3}} v_{ed} \tanh\left(\sum_{c=1}^{M_{h_2}} v_{dc} \tanh\left(\sum_{b=1}^{M_{h_1}} v_{cb} \phi\left(\sum_{a=1}^{M_x} v_{ba} x_a\right)\right)\right)\right)\right)\right). \tag{2}$$

We removed the subscript j in (2) to avoid confusion with indexing the coordinate vector $\boldsymbol{x}$ as x_a. As shown in figure 2(b), connection weights for the MPN are denoted by $\{w_{lc}, w_{cb}, w_{ba}\} \in \mathbb{R}^{M \times N}$ and weights of the SN are $\{v_{ba}, v_{cb}, v_{dc}, v_{ed}, v_{fe}, v_{if}\} \in \mathbb{R}^{M \times N}$, where the dimensions M and N are determined by the number of nodes in the hidden layers associated with a connection weight. The first and last layers of the SN use the logistic activation function $\phi(x) = 1/(1 + e^{-x})$ instead of $\tanh(x)$ because the output $\boldsymbol{S}_{\boldsymbol{x}}^{\varepsilon}$ is always positive and a negative valued coordinate input will map to a positive number. Our choice of activation functions for the SN is important since we do not implement node biases. Similar to the MPN, coordinate inputs of the SN are shifted and scaled to reside within ± 0.8. Spatial scaling values at the output of the SN are shifted and scaled to $[0.1, 0.8]$ before training.

Fixing values for the activation functions and S^{σ}, the output of the MPN for a given strain vector input can be altered by adjusting the connection weights of either the MPN or SN (the latter also changes $\boldsymbol{S}_{\boldsymbol{x}}^{\varepsilon}$). The inverse problem of estimating material properties from force-displacement measurements is often posed as an optimization problem. The parameters of a pre-selected constitutive model are identified to match u_k^n, which are displacements computed in a forward FEA, to $\hat{u}_k^n$ that are measured experimentally. We replace the constitutive model with a CaNNCM and train the network (i.e. find weights of the MPN and SN) to minimize the objective function in (3). Given N_p load increments and displacements estimated at N_d nodes in the FE mesh, the objective function minimized by training the MPN and SN in AutoP (R_m and R_s, respectively) is the L_1 norm of the difference between displacement vectors $\boldsymbol{u}_k^n$ and $\hat{\boldsymbol{u}}_k^n$:

$$R_m, R_s = \underset{R_{\theta_m}, R_{\theta_s} \in \mathbb{R}}{\operatorname{argmin}} \sum_{n=1}^{N_p} \sum_{k=1}^{N_d} |\boldsymbol{u}_k^n\{R_m, R_s\} - \hat{\boldsymbol{u}}_k^n|, \tag{3}$$

where R_{θ_m} and R_{θ_s} correspond to the weights of the MPN and SN, respectively. A thorough description of how AutoP minimizes (3) is provided in Hoerig *et al* (2019) and is summarized in figure 2(c). Here we briefly review the AutoP training process in order to introduce a new regularization term.

2.5. AutoP iterations

Two FEAs are solved in every iteration of AutoP to generate stress-strain training data from force-displacement measurements. Briefly, the stress equilibrium requirement of FEA relates forces to stresses and the compatibility requirement relates displacements to strains. In FEA_{σ}, measured forces from a given load increment are applied as BCs to the FE mesh. We can be confident the stresses computed in FEA_{σ} are physically consistent estimates of the actual stress in the material because of stress equilibrium requirements. Similarly, measured displacements are applied as constraints in FEA_{ε} and the computed strains are assumed to be physically consistent estimates of the actual strain due to compatibility requirements. The MPN is then *retrained* during the kth AutoP iteration using stresses $\hat{\sigma}_l^n(\boldsymbol{x}_j)$ computed in FEA_{σ} and strains $\hat{\varepsilon}(\boldsymbol{x}_j)$ computed in FEA_{ε} at N_x' mesh integration points[2]:

$$R_m^{k+1} = \underset{R_{\theta_m} \in \mathbb{R}}{\operatorname{argmin}} \frac{1}{2} \sum_{j=1}^{N_x'} \sum_{n=1}^{N_{\sigma}} \sum_{l=1}^{N_c} (\hat{\sigma}_l^n(\boldsymbol{x}_j) - \sigma_l^n(\hat{\varepsilon}(\boldsymbol{x}_j)); R_{\theta_m}^k, S_{\boldsymbol{x}_j}^{\varepsilon}))^2, \tag{4}$$

[2] N_x refers to the total number of integration points in the FE mesh and $N_x' \subseteq N_x$.

where a semicolon indicates the term is parameterized by the following values and N_σ refers to the number of stress-strain training pairs at $\boldsymbol{x}_j$. $N_\sigma > 1$ when implementing a training window and frame invariance (see Hoerig *et al* (2019) or appendix A). Note that (4) refers to backpropagation training.

The objective function in (4) is further minimized by adjusting the spatial scaling values $\boldsymbol{S}^\varepsilon_{\boldsymbol{x}}$,

$$\hat{S}^\varepsilon_{\boldsymbol{x}_j} = \underset{S^\varepsilon_{\boldsymbol{x}_j}\in\mathbb{R}}{\operatorname{argmin}} \frac{1}{2}\sum_{j=1}^{N_x}\sum_{n=1}^{N_\sigma}\sum_{l=1}^{N_c}(\hat{\sigma}^n_l(\boldsymbol{x}_j) - \sigma^n_l(\hat{\boldsymbol{\varepsilon}}(\boldsymbol{x}_j)); R^{k+1}_{\theta_m}, S^\varepsilon_{\boldsymbol{x}_j}))^2. \tag{5}$$

The retrained MPN R_{θ_m} is used in (5) so that any information about the spatially-varying stress-strain relationship that cannot be captured by the MPN connection weights—in this case the spatial distribution—will be captured by the spatial scaling values (and therefore the SN connection weights). An algorithm for minimizing (5) is given in Hoerig *et al* (2018). We later devised a significantly faster version of the algorithm that is presented in appendix F of Hoerig (2018). Our initial algorithm was sufficient for training 2D CaNNCMs in a reasonable amount of time. However, the same algorithm operating on the amount of data generated in 3D would require several hours to complete. Because the spatial values are recomputed during every iteration of AutoP, the creation of a faster algorithm was paramount to reducing AutoP computation time and to make training 3D CaNNCMs feasible.

After solving (5), the last two steps in AutoP are to train the SN with the newly computed spatial scaling values and check for AutoP convergence. Completion of all five steps—solve FEA_σ and FEA_ε, train MPN, compute $\boldsymbol{S}^\varepsilon_{\boldsymbol{x}}$, train SN, convergence check—constitutes one AutoP iteration. One step in AutoP is comprised of one or more iterations for force-displacement measurements corresponding to a given load increment. Cycling over all load increments/steps is one pass. Multiple passes over the force-displacement data are usually required for a CaNNCM to learn the correct material properties. In these study, 5 to 12 passes are implemented, depending on the number of data sets input to AutoP during training. Once fully trained, the weights in SN and MPN form a nonparametric model of the experimental behavior captured by the ultrasonic echo data and surface force measurements.

Further investigations with AutoP and CaNNCMs revealed shortcomings not encountered in our previous report. Specifically, material property distributions that result in regions of stress-strain concentrations are not well captured by CaNNCMs trained in AutoP. For example, the region of background material between two stiff inclusions aligned along the loading direction can appear significantly softer. We found this behavior to be caused at least in part by the inability to measure the distribution of forces at the contact surface between the probe and phantom. More fundamentally, the single measurement of axial force applied by the probe during compressive loading does not supply enough information about the distribution of stresses within the object.

2.6. Stress-matching regularization

A new regularization term introduced to (5) helps extract spatial information from measurement data. We note that both stresses and strains are computed in FEA_ε, although the former are not physically consistent with the true stresses, at least in early iterations of AutoP. Nonetheless, stresses $\bar{\sigma}^n_l(\boldsymbol{x}_j)$ computed in FEA_ε do contain a significant amount of information pertaining to the spatial distribution of stresses. We thus add a regularization term, which we refer to as a σ-matching term, that forces the spatial values to match $\hat{\sigma}^n_l(\boldsymbol{\varepsilon}(\boldsymbol{x}_j))$ and $\bar{\sigma}^n_l(\boldsymbol{\varepsilon}(\boldsymbol{x}_j))$. We first compute the magnitude of the stress differences:

$$\Delta\sigma^n_l(\boldsymbol{x}_j) = \left|\hat{\sigma}^n_l(\boldsymbol{x}_j) - \bar{\sigma}^n_l(\boldsymbol{x}_j)\right| \cdot \operatorname{sign}\left(\hat{\sigma}^n_l(\boldsymbol{x}_j)\right). \tag{6}$$

We then adjust the values in (6) to have zero spatial mean in each component of stress before adding the σ-matching term to (5):

$$\mu^n_l = \frac{1}{N_x}\sum_{j=1}^{N_x}\Delta\sigma^n_l(\boldsymbol{x}_j), \tag{7}$$

$$\mathcal{R}(\Delta\sigma^n_l(\boldsymbol{x}_j)) = \Delta\sigma^n_l(\boldsymbol{x}_j) - \mu^n_l, \tag{8}$$

$$\hat{S}^\varepsilon_{\boldsymbol{x}_j} = \underset{S^\varepsilon_{\boldsymbol{x}_j}\in\mathbb{R}}{\operatorname{argmin}} \frac{1}{2}\sum_{j=1}^{N_x}\sum_{n=1}^{N_\sigma}\sum_{l=1}^{N_c}\left(\hat{\sigma}^n_l(\boldsymbol{x}_j) - \sigma^n_l\left(\hat{\boldsymbol{\varepsilon}}(\boldsymbol{x}_j); R^{k+1}_{\theta_m}, S^\varepsilon_{\boldsymbol{x}_j}\right) + \beta_\sigma \overbrace{\mathcal{R}(\hat{\sigma}^n_l(\boldsymbol{x}_j), \bar{\sigma}^n_l(\boldsymbol{x}_j))}^{\sigma-\text{matching}}\right)^2, \tag{9}$$

Table 2. Description of the AutoP analyses. Data sets are summarized in table 1. The addition of 'L' to a data set number signifies large (7%) strain applied by compressor.

Test #	Data Sets	Passes	P_β/β_σ	η_σ	N_σ	N_p	N_d
1	S1′	5	0	0.5	150	4	19113
2	S9′					4	
3	S1′		2/1			4	
4	S9′					4	
5	S1	5	2/1	0.5	150	4	210
6	S9					4	
7	S1, S2, S3	7	4/1		400	12	630
8	S9, S10, S11					12	
9	1	5	2/0.5	0.5	150	4	225
10	9		2/1			4	
11	4		2/0.5			4	
12	16		2/1			4	
13	1, 8	5	3/0.5	0.5	250	8	450
14	9, 16		3/1			8	
15	1, 2, 3	7	4/0.5		400	12	675
16	9, 10, 11		4/1			12	
17	1 L	5	2/0.5	0.2	200	6	225
18	9 L		2/1	0.2	200	6	
19	8 L		2/0.5	0.1	250	6	
20	16 L		2/1	0.1	250	6	
21	1 L, 8 L	5	3/0.5	0.1	450	12	450
22	9 L, 16 L		3/1	0.1		12	
23	1 L, 2 L, 3 L	7	4/0.5	0.2	600	18	675
24	9 L, 10 L, 11 L		4/1	0.2		18	
25	1–7	12	8/0.5	0.5	800	28	1575
26	9–15		8/1			28	
27	1 L–7 L		8/0.5			42	
28	9 L–15 L		8/1			42	

where μ_l^n in (7) is the spatial mean stress and $\mathcal{R}(\cdot)$ in (8) is the σ-matching term. Recall that N_x refers to the total number of integration points in the FE mesh. The free parameter β_σ is a constant scalar that controls the influence of σ-matching regularization. Much like the decision to use the L_1 norm in the objective function of (3), the absolute difference is utilized in the σ-matching term of (6) to mitigate the effects of large outliers. Furthermore, neither stress value is explicitly dependent on $\boldsymbol{S}_{\boldsymbol{x}}^{\varepsilon}$ and the term in (8) is constant, meaning it has no effect on the gradient during the optimization procedure. Hence why the σ-matching term is included within the L_2 norm of (9). Finally, we add the $\text{sign}(\hat{\sigma}_l^n(\boldsymbol{x}_j))$ component because the absolute value is always positive.

Adjusting the σ-matching term in (8) helps ensure the spatial mean magnitude of $\hat{\sigma}_l^n(\boldsymbol{\varepsilon}(\boldsymbol{x}_j))$ does not change. A shift in the mean stress would alter the stress-strain relationship, causing the CaNNCM to learn incorrect material properties. Preliminary studies revealed that we could skip the adjustment in (8) and still recover the Young's modulus from the CaNNCM if β_σ was reduced. However, CaNNCMs trained in this manner typically required additional passes in AutoP and thus longer training time. Interestingly, our preliminary investigations also revealed that it is not necessary to impose σ-matching for the entire duration of AutoP training. Several different methods of choosing β_σ were explored, including more sophisticated approaches similar to learning rate scheduling used in deep learning. We found that setting $\beta_\sigma > 0$ for the first P_β passes in AutoP and then reducing it to $\beta_\sigma = 0$ is simple and effective. The choice of P_β is dependent on the number of data sets used for training and should increase with the number of separate sets of measurements input to AutoP. We also found that the value of β_σ should be reduced as the level of noise/error in the force-displacement measurements increases.

2.7. AutoP analyses

3D QUSE using data acquired with a 1D linear array probe introduces a unique sampling problem. In this study, we begin to explore sampling requirements to accurately characterize linear-elastic materials in 3D and reconstruct the Young's modulus distribution with CaNNCMs. Table 2 summarizes the AutoP parameters and data sets used to train 28 different CaNNCMs. Also included are two free parameters η_σ and N_σ used in

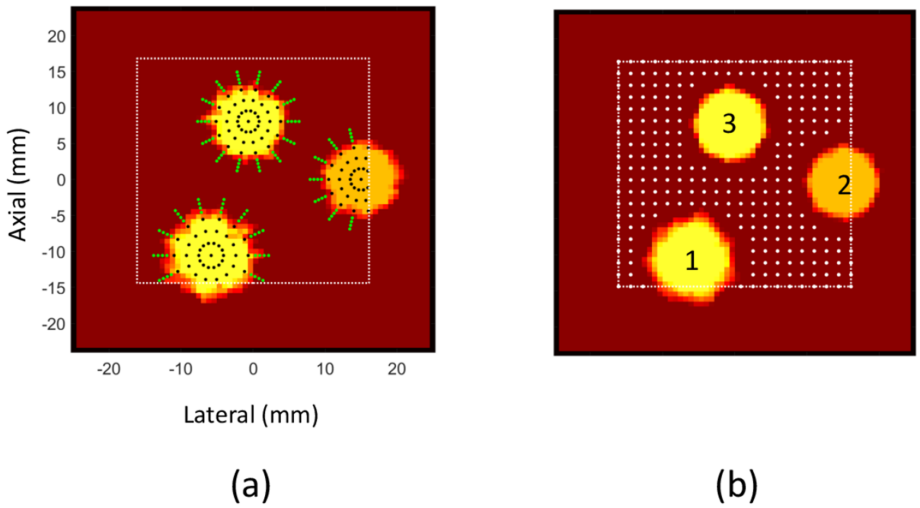

Figure 3. (a) Contrast ratios were computed by sampling the inclusions (black points) and a region of the background surrounding each inclusion (green points). Young's modulus values were estimated at each point using (11). Contrast values were calculated as the mean Young's modulus estimated at the black points for each inclusion divided by the mean Young's modulus of the green points surrounding said inclusion. (b) To obtain an estimate of the Young's modulus for the background material, we sample the space uniformly (white points) and computed the mean Young's modulus. The white bounding box indicates the image region in which interior displacements are estimated via speckle-tracking. Inclusions are labeled '1', '2', and '3' to more easily identify regions.

the algorithm to compute $\boldsymbol{S}_{\boldsymbol{x}}^{\varepsilon}$ (Hoerig *et al* 2018). The notation P_β/β_σ denotes the number of AutoP passes for which β_σ is equal to the given value. After P_β passes we set $\beta_\sigma = 0$, which removes the σ-matching term from training. Values selected for P_β were based on preliminary testing, but the general guide is to increase its value as the number of data sets input to AutoP increases. Training parameters common to all tests are described in appendix A.

After training each CaNNCM, the Young's modulus was reconstructed over the entire domain of the phantom by first defining a strain vector $\boldsymbol{\varepsilon}_{test} = [0.0035, -0.01, 0.0035, 0.001, 0.001, 0.001]$. Values for $\boldsymbol{\varepsilon}_{test}$ were chosen to represent a compressive load within the range of training data. This strain vector was then used to compute the stiffness matrix $\hat{D}_{ij}$ from the connection weights of the MPN (appendix B of Hoerig *et al* (2019)). Relating $\hat{D}_{ij}$ to the linear-elastic constitutive model and rearranging the equations, one can arrive at an expression for Poisson's ratio and Young's modulus E at location $\boldsymbol{x}$,

$$\hat{\nu}(\boldsymbol{x}) = \frac{\frac{\hat{D}_{11}}{S_{\boldsymbol{x}}^{\varepsilon_1}} + 2\frac{\hat{D}_{12}}{S_{\boldsymbol{x}}^{\varepsilon_2}}}{\frac{\hat{D}_{11}}{S_{\boldsymbol{x}}^{\varepsilon_1}} + \frac{\hat{D}_{12}}{S_{\boldsymbol{x}}^{\varepsilon_2}}} - 1 \tag{10}$$

$$\hat{E}(\boldsymbol{x}) = S^{\sigma}\left(\frac{\hat{D}_{11}}{S_{\boldsymbol{x}}^{\varepsilon_1}} - \frac{\hat{D}_{12}}{S_{\boldsymbol{x}}^{\varepsilon_2}}\right)(1 + \hat{\nu}(\boldsymbol{x}))\,. \tag{11}$$

$S_{\boldsymbol{x}}^{\varepsilon_1}$ and $S_{\boldsymbol{x}}^{\varepsilon_2}$ denote the lateral and axial components of $\boldsymbol{S}_{\boldsymbol{x}}^{\varepsilon}$, respectively, and change as $\boldsymbol{x}$ varies over the domain of the phantom.

Accuracy of the Young's modulus reconstructions were quantified two different ways. First, the contrast ratio between each inclusion and the surrounding background material was computed. Figure 3(a) depicts the sample points where $\hat{E}(\boldsymbol{x})$ was estimated using (11). The mean calculated from each set of $\hat{E}(\boldsymbol{x})$ was the estimated Young's modulus for the corresponding inclusion. Then, the mean Young's modulus was computed for a portion of the background material surrounding each inclusion (figure 3(a), green dots). The size of the surrounding strips were chosen to have approximately the same area as the related inclusion. Contrast ratios were determined as the Young's modulus of the inclusion divided by the Young's modulus of the surrounding background material. We also estimate the mean Young's modulus for the background material from points sampled in figure 3(b) (white dots). This value is later reported as the Young's modulus for the background but is not used for computing the contrast ratios. The background was sampled differently for computing contrast ratios and estimating Young's modulus because the former compares each inclusion to its immediate surrounding whereas the latter should be an average value over the scan plane.

The second metric we chose to quantify the accuracy of Young's modulus estimates was the normalized mean absolute error (NMAE). For this, mean of the estimated $\hat{E}(\boldsymbol{x})$ values for each region was compared to

the target value E_i estimated via macro-indentation measurements,

$$e_i^E = \frac{|E_i - \langle \hat{E}_i(\boldsymbol{x}) \rangle|}{E_i}, \tag{12}$$

where $i = 0, 1, 2, 3$ indicates region number (0 corresponds to background, 1–3 are the inclusions as labeled in figure 3(b)) and $\langle \cdot \rangle$ is used to indicate the mean. We report the mean and standard deviation of these four error values for each trained CaNNCM. Measuring regional moduli and their associated contrast values reveals the accuracy of the estimates as well as the relative appearance of inclusions in images.

3. Results

3.1. Simulated force-displacement measurements

Data sets S1′ and S5′ were used to train CaNNCMs in Tests 1–4. Here, we demonstrate the need for the σ-matching term by testing AutoP under nearly ideal circumstances: axial displacement measurements are available at every point in the FE mesh, not just within the scan plane beneath the US probe. Comparing the results for a large US transducer compressing the phantom (Tests 1, 3) to a small probe (Tests 2, 4) will help reveal any fundamental differences or limitations to acquiring data with a smaller probe.

CaNNCMs trained in Tests 5–8 also utilized simulated, noise-free data, albeit displacements were limited to only within the scan plane. A single data set was input to AutoP for Tests 5 and 6 whereas three data sets were input for Tests 7 and 8. Results from these analyses demonstrate the effect additional data has on the ability of CaNNCMs to accurately estimate material properties throughout a larger region of the phantom. It is expected that a CaNNCM trained with a single data set whose displacements are constrained to a plane will be able to accurately estimate both material parameters and geometry only within said plane. Adding additional data sets by moving the probe to acquire data in parallel planes should provide the additional information necessary for a CaNNCM to estimate material properties throughout more of the volume.

Young's modulus images reconstructed by CaNNCMs trained in Tests 1–8 using (10) and (11) are displayed in figure 4. We show the image Planes 1–3 and a surface plot the Young's modulus, both described in figures 1(c) and (d), respectively. A surface plot of the target distribution, shown as a translucent gray surface, is included to enable direct comparison with the CaNNCM estimates. Young's modulus and Poisson's ratio estimates for the phantom material, contrast ratios, and mean absolute error as computed with (12) are compiled in table 3.

Comparing Tests 3 and 4 where $\beta_\sigma = 1$ with Tests 1 and 2 where $\beta_\sigma = 0$ clearly demonstrates an improvement gained by adding σ-matching. Without this term, the spaces between and around the inclusions coinciding with regions of high stresses and strains appear too soft. In turn, the estimated modulus values for the inclusions are too low.

Reducing the amount of available displacement data to only the image plane has a significant effect, as demonstrated by the results for Tests 5–8. We observe that, within the image plane, the CaNNCM accurately estimates the inclusion shapes and the Young's modulus of the background (Tests 5 and 6) but estimates of the inclusion stiffness are too large. Including additional data sets during training with interior displacement data in adjacent image planes not only extends the volume over which the CaNNCM learns the inclusion geometries (Tests 7 and 8), but also generally improves the Young's modulus estimates.

These results reveal a limitation of the method when using a smaller compression surface: limited imaging depth. It is already known that adding a force plate to the US probe for QUSE increases the possible image depth. Our results confirm this limitation for the current formulation of AutoP and CaNNCMs. The limited imaging depth is a result of the decreased deformation far away from where the compressive load is being applied. Even for cases like Test 4 where noise-free axial displacements are known at all points in the mesh, away from the US probe the Young's modulus reconstructed by the CaNNCM contains errors in the estimated inclusion shape. As the available displacement data is reduced to the RF echo scan plane (Tests 6 and 8), the CaNNCM captures the shape of Inclusion 1 but significantly underestimates the Young's modulus.

3.2. Experimental force-displacement measurements

All remaining tests were performed using force-displacement measurements acquired experimentally on a gelatin phantom. Tests 9–12 each accept a single data set as input to AutoP for a large/small probe compressing a phantom from one of two sides (0° or 90° phantom rotation). Then, multiple data sets are used for training in Tests 13–16. Compared to the preceding tests, the use of experimentally acquired data introduces noise and measurement uncertainty.

Young's modulus reconstructions in the bottom half of figure 4 were produced by CaNNCMs trained with force-displacement measurements acquired experimentally on a gelatin phantom. CaNNCMs in Tests 9,

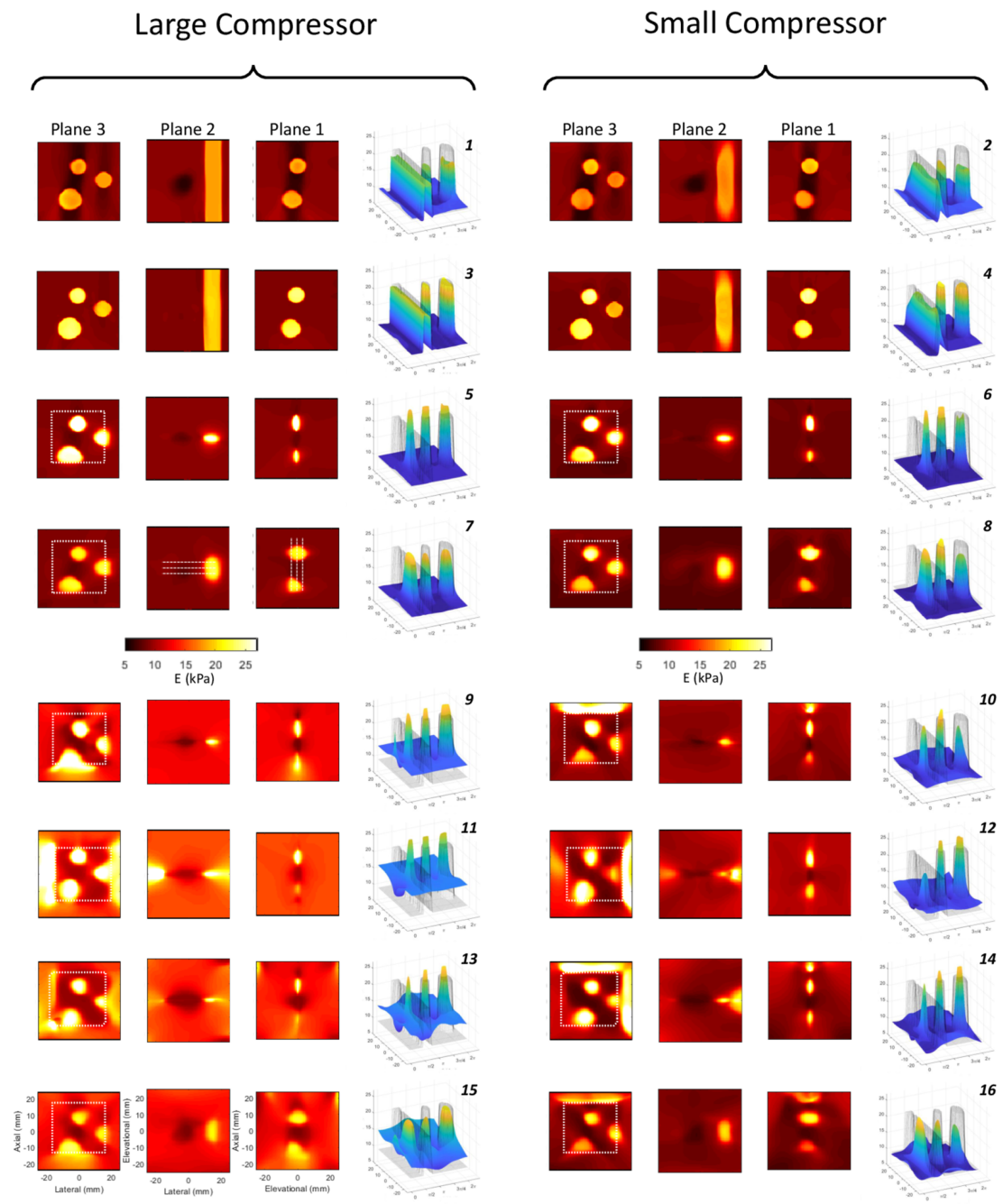

Figure 4. Young's modulus images reconstructed by CaNNCMs trained in Tests 1–16. Images in the left column are results from tests where a large compressor was used whereas the right column corresponds to the analogous tests using a small compressor. The number in the top-right corner of each image set identifies the test number from table 2. Explanations of Planes 1–3 and the Young's modulus surface plot are in figures 1(c) and (d), respectively. The white box in Plane 3 indicates the boundary of the scan plane. The dotted white lines in Planes 1 and 2 of Test 7 denote the locations of all three scan planes.

10, 15, and 16 were trained using nearly identical parameters (difference in β_σ) as those trained in Tests 9, 10, 11, and 12, except the measurement data input to the former contain echo noise and force and displacement errors. Nonetheless, similar characteristics are observed. Training in AutoP using data from a single measurement plane, as done for Tests 9 and 10, is insufficient for a CaNNCM to learn the 3D shape. Including data from other loading locations and image planes (Tests 15 and 16) increases the volume over which the CaNNCM accurately estimates material properties and geometry. Directly comparing Young's modulus reconstructions from CaNNCMs trained with simulated versus experimentally acquired data, we note the latter significantly overestimates the Young's modulus of the background outside of the scan plane, especially for when a large compressor is used. This may be partly attributed to errors in force measurements and is studied in section 3.3.

Table 3. Quantitative estimates of Young's modulus values, Poisson's ratio, contrast ratios, and mean absolute error (NMAE) for all CaNNCMs. Also included are AutoP training times. 'Target' values are those estimated via macro-indentation methods. Inc. 1–3 refer to the inclusions as identified in figure 3. Values in bold are those closest to the target value and only consider CaNNCMs trained with measurement data acquired experimentally. Spatial standard deviation of the Poisson's ratio estimated at multiple points using (10) is at least two orders of magnitude less than the spatial mean and therefore not included.

	E Estimate (kPa)				Contrast Ratio					
Test #	Back.	Inc.1	Inc. 2	Inc. 3	Inc. 1	Inc. 2	Inc. 3	ν Estimate	NMAE	
Target	8.92	24.17	20.53	24.17	2.71	2.30	2.71	0.3	$\mu\pm\sigma$	Training Time (min.)
1	8.6 ± 0.9	18.4 ± 0.7	17.9 ± 1.0	18.0 ± 0.7	2.14	2.09	2.09	0.28	0.16 ± 0.10	160
2	9.0 ± 0.1	23.5 ± 0.5	19.8 ± 1.0	23.7 ± 0.8	2.61	2.20	2.64	0.29	0.02 ± 0.01	160
3	8.4 ± 0.8	17.5 ± 0.8	17.0 ± 0.9	19.3 ± 0.6	2.08	2.02	2.30	0.27	0.18 ± 0.09	145
4	8.8 ± 0.2	23.1 ± 0.7	18.8 ± 1.0	24.4 ± 1.4	2.63	2.14	2.78	0.28	0.04 ± 0.03	145
5	9.2 ± 0.5	25.8 ± 1.5	26.0 ± 2.0	28.5 ± 1.2	2.79	2.82	3.09	0.30	0.14 ± 0.11	151
6	8.7 ± 0.7	22.0 ± 2.5	25.4 ± 2.8	28.0 ± 2.3	2.52	2.91	3.21	0.31	0.13 ± 0.09	135
7	9.3 ± 0.5	20.8 ± 2.0	22.1 ± 1.7	20.9 ± 1.5	2.24	2.38	2.26	0.30	0.10 ± 0.05	710
8	8.8 ± 0.6	18.9 ± 2.1	22.5 ± 2.1	24.1 ± 2.5	2.15	2.55	2.73	0.32	0.08 ± 0.10	659
Target								0.5		
9	11.4 ± 1.6	26.5 ± 3.6	25.3 ± 2.3	27.5 ± 1.8	2.32	**2.21**	2.41	0.32	0.19 ± 0.08	177
10	9.8 ± 2.9	18.6 ± 2.5	20.9 ± 2.9	26.1 ± 1.7	1.91	2.14	**2.67**	0.39	0.10 ± 0.09	141
11	13.1 ± 3.2	28.1 ± 4.5	27.3 ± 4.3	30.4 ± 3.2	2.15	2.08	2.32	0.33	0.31 ± 0.13	151
12	10.5 ± 3.7	25.7 ± 5.3	17.4 ± 2.6	25.1 ± 3.8	2.45	1.65	2.39	0.39	0.11 ± 0.07	142
13	11.7 ± 3.1	25.1 ± 4.6	22.2 ± 4.0	26.1 ± 2.9	2.15	1.90	2.23	0.32	0.13 ± 0.12	459
14	$\mathbf{8.8 \pm 1.3}$	$\mathbf{24.6 \pm 5.1}$	18.6 ± 3.5	25.3 ± 2.9	**2.81**	2.13	2.89	0.37	$\mathbf{0.04 \pm 0.04}$	285
15	12.2 ± 1.3	21.9 ± 2.6	$\mathbf{20.4 \pm 1.5}$	20.8 ± 1.6	1.80	1.68	1.71	0.34	0.15 ± 0.15	718
16	9.4 ± 2.4	16.9 ± 2.1	20.2 ± 2.9	20.4 ± 2.9	1.79	2.13	2.16	0.39	0.13 ± 0.13	675
17	11.1 ± 1.6	23.7 ± 3.5	22.8 ± 2.4	$\mathbf{23.8 \pm 2.5}$	2.13	2.05	2.14	0.34	0.10 ± 0.11	223
18	9.8 ± 3.0	21.6 ± 2.6	20.8 ± 3.8	22.1 ± 3.9	2.21	2.13	2.26	0.38	0.07 ± 0.04	239
19	13.4 ± 4.1	26.9 ± 4.6	24.1 ± 3.0	27.4 ± 2.9	2.00	1.79	2.04	0.36	0.23 ± 0.18	210
20	10.5 ± 4.2	30.1 ± 5.3	19.7 ± 2.5	32.8 ± 5.3	2.86	1.87	3.12	0.41	0.21 ± 0.13	251
21	11.1 ± 3.1	22.0 ± 2.9	17.7 ± 2.0	19.8 ± 1.8	1.99	1.60	1.79	0.33	0.16 ± 0.06	412
22	9.4 ± 1.9	22.3 ± 3.7	17.8 ± 3.1	23.7 ± 3.1	2.37	1.89	2.51	0.37	0.07 ± 0.05	540
23	11.6 ± 1.5	21.6 ± 3.1	20.2 ± 2.3	20.2 ± 2.3	1.86	1.74	1.74	0.35	0.15 ± 0.12	899
24	9.3 ± 2.4	18.2 ± 2.8	19.5 ± 3.8	20.8 ± 4.9	1.96	2.10	2.25	0.38	0.12 ± 0.10	1258
25	12.1 ± 1.3	18.9 ± 1.8	19.0 ± 2.5	18.5 ± 1.7	1.56	1.57	1.53	0.36	0.22 ± 0.12	1759
26	8.5 ± 2.0	12.1 ± 1.7	16.5 ± 2.9	17.5 ± 3.4	1.42	1.95	2.06	0.40	0.26 ± 0.19	1824
27	11.6 ± 1.6	18.7 ± 2.3	17.0 ± 2.5	17.1 ± 2.5	1.61	1.46	1.47	0.34	0.25 ± 0.06	3527
28	8.4 ± 2.2	11.7 ± 2.3	14.6 ± 2.7	16.0 ± 4.5	1.39	1.74	1.90	0.39	0.30 ± 0.19	3184

For all CaNNCMs trained with experimental measurements, the Young's modulus reconstruction has bands just above and below the image region where the background appears much stiffer. This artifact also appears in 2D. Biased errors in the internal displacement estimates limits the compression of the region between the scan plane and compressor, decreasing the strain in FEA_ε and thus appearing stiffer. Conversely, the band below the image region is largely the result of a mismatch in the pre-training Poisson's ratio and compressibility of the actual material. Recall that we chose $\nu = 0.3$ to generate pre-training data using (A1) to avoid numerical instabilities. We tested the effect of Poisson's ratio mismatch separately by generating a data set like S1 with the exception that $\nu = 0.495$. Using this data set and training a CaNNCM using the same parameters as Test 1, the Young's modulus image (not shown) displays a background modulus that is greater than the true value just below the image region but no significant artifacts above.

Tests 11–14 do not have simulated counterparts. Force-displacement measurements acquired after rotating the phantom 90° provides new information and measurement errors uncorrelated with prior data sets. As such, results for Tests 11 and 12 (after 90° rotation) have notable differences when compared to Tests 9 and 10 (0° rotation). First, even when compressing with the small probe, the CaNNCM is able to fully capture the geometry of each inclusion that resides within the image region. Compare the estimated shape of Inclusion 1 in Tests 10 and 12. In the latter, the shape of the inclusion is clearly circular. Second, the relative stiffness of each inclusion is preserved in Test 12 even though the Young's modulus magnitudes are fairly inaccurate. Training with data sets acquired on both sides of the phantom improved the accuracy of the Young's modulus estimates for the large compressor (Test 13) compared to either case where a single data set was used for training. Similarly, the NMAE for Test 14 is decreased when compared to Test 10, but slightly larger than Test 12. As we discuss later, we believe this is due to force measurement errors.

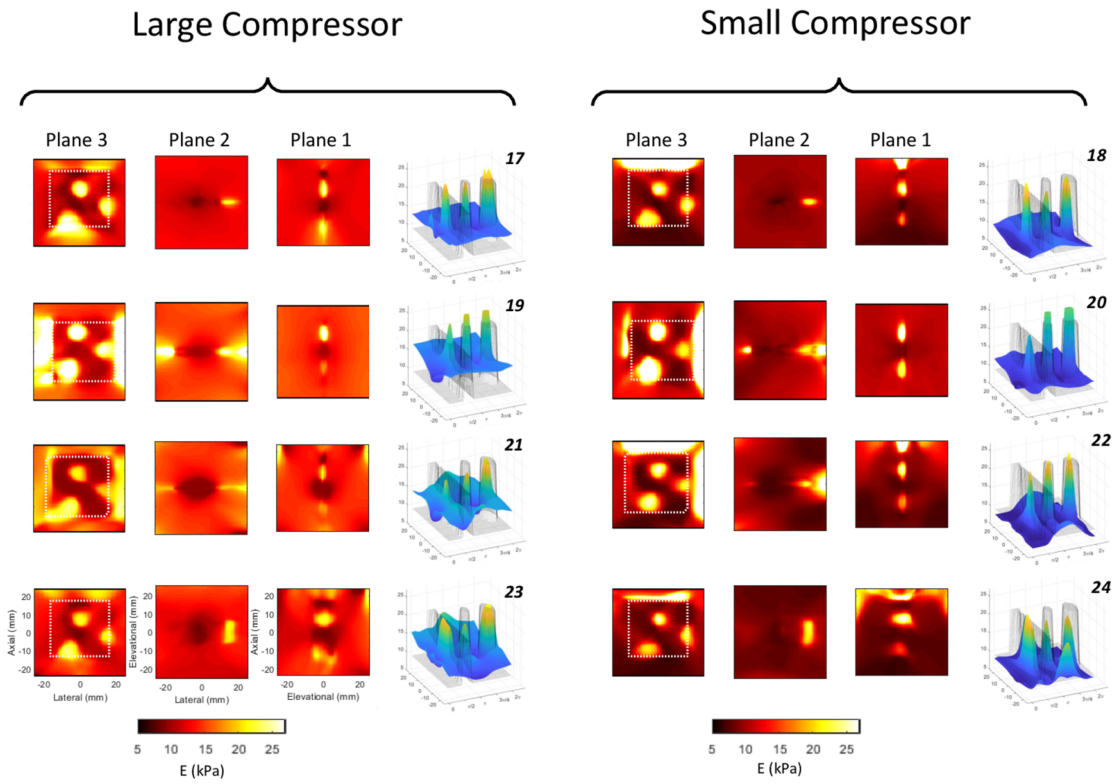

Figure 5. Young's modulus reconstructions by CaNNCMs trained in Tests 17–24 where the total compressive load induced up to 7% strain. The number in the top-right corner of each image set identifies the test number from table 2.

Tests 17–24 repeat Tests 9–16 except with 7% total applied strain as opposed to 3% (see tables 1 and 2). Results of these tests will show the effect of larger deformation on the resulting Young's modulus images. Young's modulus distributions reconstructed by CaNNCMs trained in Tests 17–24 are displayed in figure 5. In general, increasing the applied load led to reduced NMAE when measurement data were acquired with a large compressor but had no significant effect on the NMAE when data were acquired with the small compressor. However, when the small compressor applied surface loads for a phantom with 0° rotation, increasing the force load resulted in more accurate Young's modulus estimates for inclusion 1. This was expected because due to the increased deformation deeper in the phantom with the larger force load, thus increasing the SNR of the displacement estimates.

CaNNCMs in the final four tests (Tests 25–28 in table 2) were trained with measurement data acquired in seven parallel scan planes, spanning from $x_3 = -12$ to $x_3 = 12$ mm in 4 mm increments. Young's modulus images reconstructed by these four CaNNCMs are displayed in figure 6. Force-displacement data input for tests 25 and 26 were measured under 3% maximum strain with a large and small compressor, respectively. The imposed strain was 7% in tests 27 and 28. As expected, the additional measurement data provides the information necessary for the CaNNCMs to learn the internal phantom geometry over a larger region. Furthermore, data acquired under larger deformation did not greatly affect the modulus estimates, although we expected more accurate for inclusion 1 when the small compressor was used (Test 28 in figure 6).

3.3. Effect of probe force and displacement measurement errors

Results of Tests 9–24 are affected by errors in the internal displacements estimated via speckle-tracking and measurements of applied force. Given the precision of the positioning system, there is insignificant error in the exact position and orientation of the US probe/compressor over time (probe displacement errors). However, applied force and probe displacement are boundary conditions and errors affecting these measurements will likely have a global impact on training. Conversely, speckle-tracking errors are local and will only significantly affect the material properties learned in a limited region. Thus, it is important to independently study the effect of probe displacement and applied force errors on training in AutoP

The effects of probe displacement errors were tested using data set S1 and adding –0.143 to 0.143 mm offset to the total probe displacement in fourteen linearly-spaced increments. We chose a maximum magnitude of 0.143 mm error because it corresponds to 10% of the total 1.43 mm axial probe displacement. Fourteen CaNNCMs were trained in AutoP (same parameters as Test 5) using data set S1, the large

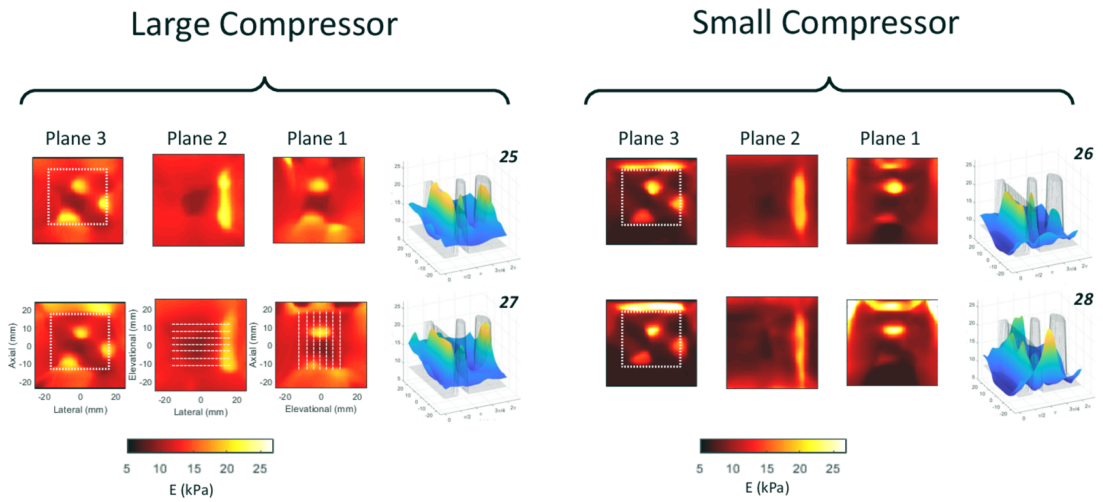

Figure 6. Young's modulus images reconstructed by CaNNCMs trained with 7 data sets. The number in the top-right corner of each image set identifies the test number from table 2. The dotted white lines in Planes 1 and 2 of Test 27 denote the locations of all seven scan planes.

compressor, and up to 10% error in the probe displacement. The tests were repeated to train another fourteen CaNNCMs with data set S9, where the small compressor is used. Similar tests were performed to investigate the influence of force measurement errors. Another fourteen CaNNCMs were trained for both large and small compressors in AutoP again using data sets S1 and S9; however, –70 mN to 70 mN (in fourteen linearly-spaced increments) was added to the applied force. Because the force necessary to induce 3% total strain is different for large and small compressors, the same magnitude force offset added to data set S9 is a larger relative error compared to the offset added to S1. We chose to investigate the effects of force measurement errors in this way because the magnitude of the error observed experimentally is systematic and not tied to the measured value.

Force errors have a fairly straight-forward effect, as shown by figure 7. Errors that increase the magnitude of the applied force input to AutoP for the same probe displacement cause the CaNNCM to overestimate Young's modulus. Conversely, modulus values are underestimated when errors reduce the magnitude of the measured force value input to AutoP. Probe displacement errors appear to have a more complex effect. For small displacement errors but no force measurement error, the resulting Young's modulus estimates are affected in a predictable way: when the true probe displacement is under-reported during training, the Young's modulus is overestimated whereas the modulus is underestimated when the reported displacement is larger than the true displacement. It appears that as the error magnitude surpasses ≈5% fractional error, the geometry of the inclusions estimated by the CaNNCM begins to distort, especially for inclusion 3. This behavior is likely due to the extreme stiffening or softening of the area between the compressor and scan plane that causes a sharp transition affecting the ability of the SN to learn the material property distribution for the given training parameters. Errors that increase the compressive load (positive abscissa in figures 7(b) and 7(c)) tend to increase the NMAE compared to errors that decrease total compression, as evidenced by the asymmetry of the curves in the aforementioned figures. Considering the phantom geometry and BCs, notably the no-slip bottom boundary, increasing the force or displacement compressive load too much may elicit geometrically non-linear effects, leading to the observed asymmetry.

The black curve in figure 7(c) begins to decrease slightly after the force error surpasses ≈+6%. The Young's modulus images corresponding to these three points are the +50 mN, +60 mN, and +70 mN labeled images in the top portion of figure 7(a). Artifacts in regions coinciding with stress-strain concentrations become increasingly prominent with the force error. These artifacts, as discussed before, cause a 'softening' of both the background and inclusions. The effect is to rescale the modulus estimates closer to the true values given that the CaNNCM is generally over-estimating the parameter values. As a consequence, the contrast ratios decrease, as shown in the top-right plot of figure 7(d). The data in figure 7 provide some insights into the effects on the accuracy and precision of modulus estimation from errors in probe position and applied force.

4. Discussion

This paper extends our data-driven approach to quantitative QUSE from 2D to 3D applications. Minor adjustments to the CaNNCM architecture enable the network to learn the heterogeneous stress-strain

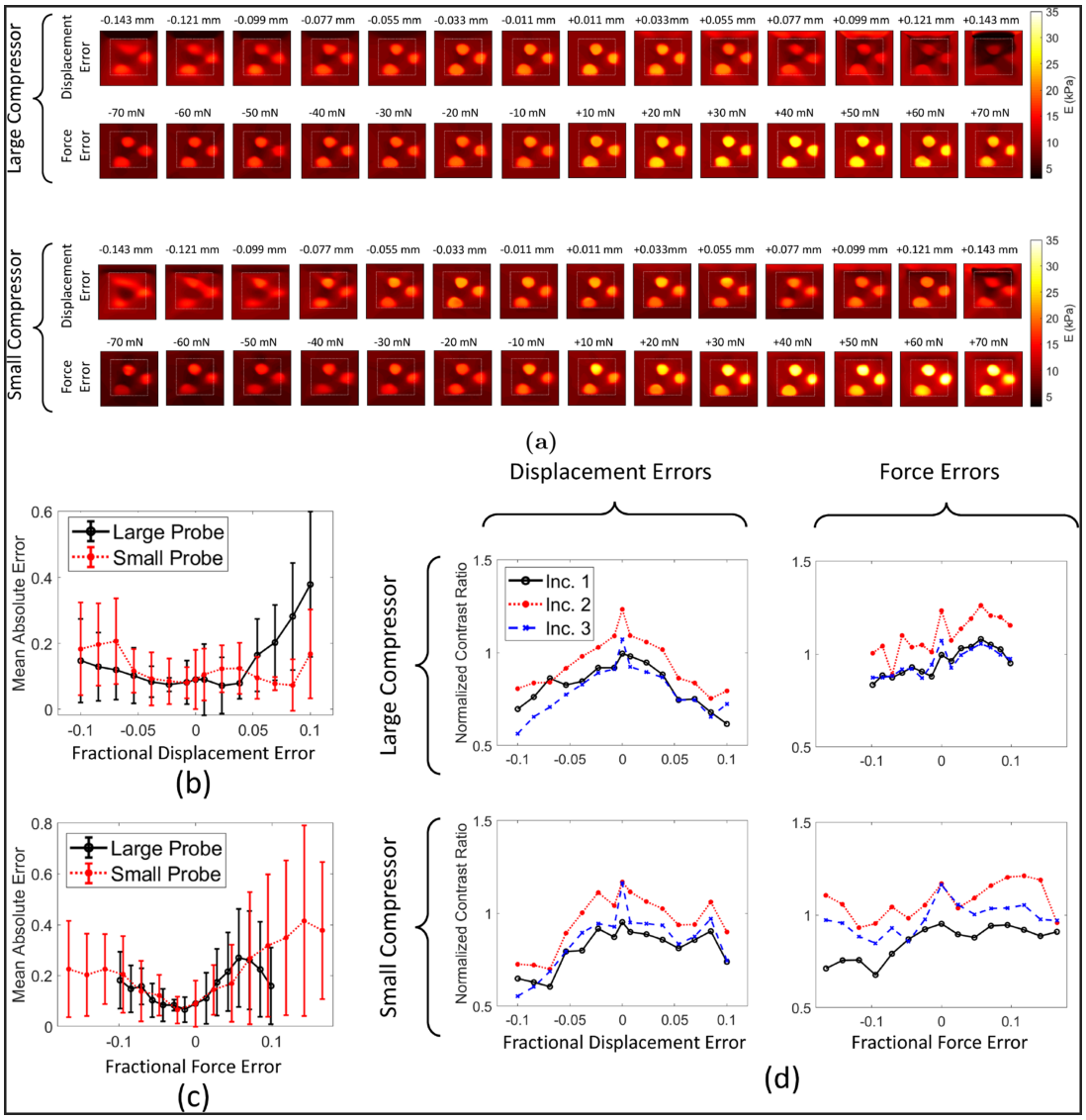

Figure 7. (a) Young's modulus images reconstructed by CaNNCMs trained with data sets S1 (top group) and S5 (bottom group). The top and bottom rows of each group corresponds to added probe displacement and force error, respectively. The values above each image indicate the error value added to the noise-free data. (b) Mean and standard deviation the Young's modulus reconstruction error computed in (12) as a function of fractional probe displacement error for both large and small compressors. (c) Mean and standard deviation of Young's modulus reconstruction error as a function of fractional force error for both size compressors. (d) Normalized contrast ratios for all three inclusions plotted against force and displacement error. Values are normalized by the inherent contrast specified as 'Target' in table 3.

relationship of material throughout a volume. The fundamentally different approach of AutoP, which combines machine learning and computational mechanics to extract stress-strain information from force-displacement data, is able to accurately estimate the Young's modulus distribution from a sparse sampling of independent measurements. This could have an advantage in clinical applications by allowing an operator to strategically select sampling locations and reduce the effects of motion artifacts.

Preliminary investigations revealed our previous implementation of AutoP is not sufficient for 3D imaging. A single measurement of surface force does not provide sufficient information about the stress distribution. Even for the best case where all axial displacements are known for a phantom under compression (Tests 1 and 2, figure 4), artifacts appear in the Young's modulus reconstruction at locations where stresses and strains are significantly increased, notably in regions between inclusions the material appears softer than it is. The effect propagates and causes the inclusions to also appear softer. One approach to correcting this shortcoming is the introduction of the σ-matching regularization term which regularizes the optimization computed by the SN to make sure the stress information is spatially consistent. Enforcing σ-matching during AutoP and returning to the case where all axial displacements are known (Tests 3 and 4), we observe significantly more accurate Young's modulus reconstructions and removal of softening artifacts in regions where stress-strain build-ups occur. Additionally, in results not shown using simulated

force-displacement data, we found that applying the known surface force distribution directly to the phantom surface (and removing the US probe) did not have a significant effect on the result. This implies that the spatial variations of force for a single compressive load do not supply sufficient information to infer the spatial distribution of stresses, meaning σ-matching is still required.

It is possible that a different regularizer could be used in place of our proposed σ-matching term and achieve the same results. Unlike other types of regularization, σ-matching relies on a fundamental principle of mechanics: imposing force BCs or the equivalent displacement BCs will result in the same distribution of stresses and strains. σ-matching exploits this principle to extract information about stress distributions missing from the force measurements.

When interior displacement measurements are limited to the scan plane, the case for Tests 5–6, the CaNNCM is able to learn the geometry and material properties of the inclusions only within the scan plane. Interestingly, reducing the displacement data to the scan plane causes the CaNNCM to overestimate the Young's modulus of the inclusions but not the background. Adding force-displacement data from additional scan planes (Tests 7–8) appears to rescale the Young's modulus estimates of the inclusions closer to the true values. Comparing images in figure 4 for Tests 6 and 8 where a small compressor was used to Tests 5 and 7, we note that using the smaller compressor led to more accurate Young's modulus estimates for inclusions 2 and 3 and a reduction in the background artifact, but the deeper inclusion 1 is more obscured. An analogous effect is seen in Test 4 where the ends of inclusion 2 far away from the compressor are not well defined. This is the result of decreased deformation away from the compressor as its size decreases.

Additional data sets generally improved the Young's modulus estimates and increased the volume over which a CaNNCM learns the internal phantom geometry at the cost of increased computational load. All CaNNCMs in this study were trained on one of two standard desktop computers, albeit one computer contained a NVIDIA TitanX GPU. At this time, the GPU was used only during the calculation of $\mathbf{S}_{\boldsymbol{x}}^{\varepsilon}$ and significantly reduced the solve time. Thus, Tests 25–28 (7 data sets) were run on the computer containing the GPU. Training time was roughly proportional to the number of data sets. However, the required time to train a CaNNCM with even a single data set makes clinical imaging infeasible. We expect to reduce training time by running all FEA_{σ} and FEA_{ε} for each AutoP iteration in parallel and taking advantage of GPU resources for solving FEAs and network training. These tasks will require the development of a custom FEA solver.

We found that care must be taken when choosing β_{σ} and the number of passes for which σ-matching is enforced. For reference, $0 \leq \beta_{\sigma} \leq 1$, where $\beta_{\sigma} = 0$ corresponds to no matching and $\beta_{\sigma} = 1$ indicates complete matching. Because the gradient of (9) is not affected by the σ-matching term, selecting a large value for β_{σ} can cause a positive feedback effect and lead to severe overcorrection of the stresses. On the other hand, setting β_{σ} too small requires more AutoP passes in order to be effective, increasing the total training time. Preliminary results suggested β_{σ} is sensitive to measurement noise and modeling errors. Therefore, we chose $\beta_{\sigma} = 0.5$ for CaNNCMs trained with experimental measurement data acquired with a large compressor. Tests not shown revealed that increasing the value of β_{σ} in these cases would result in significantly larger NMAE due to the CaNNCM overestimating Young's values.

Interestingly, parameter estimates were most accurate when $\beta_{\sigma} = 1$ for all CaNNCMs trained with simulated, noise-free data sets and experimental measurements acquired using a small compressor. We note that estimates of the background material are more accurate in locations away from the scan plane(s) when training with data acquired using a small compressor. To the best of our knowledge, this is a result of incorrect FE modeling of the contact area between the large compressor and phantom for a pre-strained state. We expect similar levels of noise in the experimentally acquired force-displacement measurements when using a large or small compressor, leaving FE modeling as a variable source of error. Unfortunately, we are unable to accurately measure the pre-strained phantom geometry, nor the contact area between the phantom and large compressor. In such cases it is therefore advantageous to acquire data with a small compressor due to the unchanging contact conditions.

Force-displacement measurements acquired experimentally have errors that impact the material properties learned by a CaNNCM in different ways. We previously showed that noise added to the internal displacement estimates can distort the estimated geometry and result in the band of increased stiffness just above the scan plane. Comparing Young's modulus reconstructions by CaNNCMs from Tests 9–10 with their counterparts trained with simulated data (Tests 5–6), we note increased artifacts and erroneous stiffening of the background material, particularly in the spaces between inclusions and edges of the scan plane (Hoerig *et al* 2019). Rotating the phantom and acquiring force-displacement measurements from a different side results in Young's modulus reconstructions (Tests 11, 12, 19, 20) with different artifacts. Training a CaNNCM with both data sets, as done in Tests 13, 14, 21, and 22, appears to reduce the influence of noise, implying the errors were uncorrelated between data sets. More fundamentally, rotating the phantom and applying compressive loads on a different surface represents a completely different set of BCs that increases the variation in the total set of stress-strain data. As a consequence, the CaNNCMs trained with data sets

acquired under completely different loading conditions tend to more accurately learn the linear-elastic properties of the gelatin phantom. This finding emphasizes the need for diversity in the training data.

The estimated Poisson's ratio in Tests 9–28 never approached the expected value of 0.5. We do observe, though, that the Poisson's ratio was closest to the expected value in tests where a small compressor was used to compress the phantom in several locations. These results suggest the CaNNCMs are able to learn the *shear* modulus from the available measurement data but have difficulty extracting the *bulk* modulus, yet estimating the Poisson's ratio from the CaNNCM in (10) instead of imposing an incompressibility constraint ($\nu = 0.5$) for the gelatin phantom still allows for accurate estimation of the Young's modulus. Additional studies will be necessary to more fully investigate how different sampling strategies affect the learned shear and bulk material properties.

Errors in the probe force and displacement measurements can have a significant effect on parameter estimates because they are boundary conditions. Such errors propagate through to the stresses and strains computed in FEA_σ and FEA_ε. Therefore, the 'reference' material learned by a MPN will be most affected by errors in the measured applied force and probe displacement. Because of the high precision of the positioning system, we can be confident that Young's modulus estimation errors for CaNNCMs trained with experimental measurements are predominantly caused by errors in FE model, measured force, and local displacements estimated via speckle-tracking. Nonetheless, these results demonstrate the importance of accurate measurements of applied force and probe position during data acquisition.

Increasing the compressive load may increase the SNR of both the force measurements and internal displacement estimates. As a consequence, the effect of systematic noise is reduced and the contrast increased in the Young's modulus image. Comparing Young's modulus reconstructions in figure 5 to their counterparts in figure 4 shows a reduction in the background artifacts and decrease in NMAE. Interestingly, the effect is most pronounced when a large compressor was used. Results in figure 7 do not suggest data acquired with a small compressor are more robust to noise than data acquired with a large compressor. Rather, it may be the boundary conditions are better modeled in the FEA for a small compressor. As previously discussed, the FE mesh was created to model the phantom in its pre-strained state. Whether the top surface of the phantom was 50×50 mm^2 (no pre-strain) or 51×51 mm^2 (with pre-strain), the contact area between the phantom and small compressor does not change. On the other hand, the contact area between the phantom and large compressor will change based on mesh size and grow as the applied load increases. This is an example of a non-linear effect not currently accounted for in AutoP.

We also observe increased distortion of the internal geometry estimated by CaNNCMs when trained with data acquired under larger compression. This is evident when comparing Plane 1 of Tests 26 and 28 in figure 6. Specifically, the shape of Inclusion 3 resembles a sphere under the small strain case but appears oblong when the CaNNCM was trained with large deformation data. To the best of our knowledge, this is a result of 1) summing multiple small displacement estimates to obtain an estimate of the total displacement and 2) small-strain FEA instead of a geometrically non-linear FEA formulation. The former is required to avoid decorrelation between RF echo frames and obtain valid displacement estimates with speckle-tracking. It may be possible to avoid summing small displacement estimates by utilizing an updated Lagrangian FEA formulation and imposing incremental displacement BCs instead of the total displacement Ghaboussi *et al* (2017).

Despite the potential benefits, acquiring force-displacement data under large compressive loads will almost surely induce a non-linear elastic material response from tissues and activate geometrically non-linear effects. Previous studies using AutoP have shown the capabilities of neural network constitutive models to model highly non-linear behaviors under small deformation (e.g. Hashash *et al* (2006a), Hashash *et al* (2006b), Jung and Ghaboussi (2006), Jung and Ghaboussi (2010), Kim *et al* (2012)). Our future work with CaNNCMs will aim to combine large deformation and material non-linearity to effectively capture the mechanical properties of soft tissues.

5. Conclusions

3D Cartesian neural network constitutive models trained in the Autoprogressive Method can learn the material properties of a volume from force-displacement measurements acquired in a sparse sampling of scan planes. The previous implementation of AutoP is insufficient for CaNNCMs to accurately learn both material properties and interior geometry. A new regularization term specific to AutoP and referred to as σ-matching effectively extracts stress distributions from force-displacement measurements by coupling spatial information encoded within displacement data to stress equilibrium requirements inherent to finite element analysis. The ability of CaNNCMs to learn spatial material property distributions is further improved when measurement data are diverse, which can be achieved by reducing the compressor size and applying loads at a variety of surface locations. The method is currently only capable of characterizing

linear-elastic mechanical properties and must be extended to non-linear, hyperelastic, and time-dependent materials for effective clinical imaging.

Acknowledgments

Research reported in this publication was supported by NCI and NIBIB of the National Institutes of Health under Award Numbers R01 CA168575 and R21 EB023402. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. We gratefully acknowledge the support of NVIDIA Corporation with the donation of Titan X and Titan Xp GPUs used in this research. The authors would also like to thank H. Hashemi and H. Rivaz for providing the source code for their displacement estimation algorithm.

Appendix A. AutoP training parameters

A.1. Network specifications

Compared to 2D applications, three changes were made to the network architecture: 1) the number of input and output nodes of the MPN increased from three to six; 2) the number of input and output nodes of the SN increased to three and six, respectively; and 3) increased the number of hidden to layer nodes to 10 and 20 for the MPN and SN, respectively ($N_{h_1} = N_{h_2} = 10$ and $M_{h_1} = M_{h_2} = M_{h_3} = M_{h_4} = M_{h_5} = 20$ in figure 2(b)).

In this study, the MPN had two hidden layers and all nodes used a hyperbolic tangent activation function with no node biases. The last two points ensure that a zero-valued strain input produces a zero-valued stress. Keeping the number of hidden layers small helps avoid overtraining The network was trained using a custom implementation of the iRPROP+ algorithm Riedmiller and Braun (1993) for 15–50 epochs each AutoP iteration. Because the hyperbolic tangent function saturates as the input increases beyond ± 1, stresses computed at the MPN output are scaled by S^{σ} to ensure they remain within a range of ± 0.8. S^{σ} is checked every iteration of AutoP but typically does not deviate from the value chosen during pre-training. Conversely, the SN contained five hidden layers. A logistic activation function was used for the first and last layers and a hyperbolic tangent was used for all other layers, as specified in (2). We chose to use a logistic activation function at the input and output layers because a zero-valued coordinate input should not map to a zero-valued output and the output is always positive. More modern activation functions like ReLU/ELU were tried in preliminary testing, but we were unable to achieve satisfactory results. Inputs and outputs to the SN were shifted and scaled to within ± 0.8 to avoid saturating the nodes. Full batch training of the SN was implemented in TensorFlow using the Adam optimizer (with default settings) and a learning rate of 0.01 for 1000 epochs.

A four-load training window was implemented which includes the stress-strain data from the prior four AutoP iterations during MPN retraining and $\boldsymbol{S}_{\boldsymbol{x}}^{\varepsilon}$ calculation in the current iteration. Furthermore, we enforced frame invariance by rotating the original stress-strain data (computed in FEA_{σ} and $\mathrm{FEA}_{\varepsilon}$) around all three axes, effectively quadrupling the total amount of training data each iteration. We reduced the number of data pairs used for MPN retraining by randomly selecting 400 elements each iteration and collecting the stress-strain data from those elements only, leaving a maximum of $400 \times 4 \times 4 = 6400$ training pairs. However, all $56172 \times 4 \times 4 = 898752$ stress-strain pairs were input to (5) every iteration to recompute $\boldsymbol{S}_{\boldsymbol{x}}^{\varepsilon}$ and all 56172 $\boldsymbol{x}$–$\boldsymbol{S}_{\boldsymbol{x}}^{\varepsilon}$ pairs were used for SN training. AutoP convergence criteria were the same as specified in Hoerig *et al* (2019).

A.2. Linear-elastic Pretraining

Before training in AutoP, stress-strain data were generated using linear-elastic equations to pre-train each CaNNCM. We assign Young's modulus E and Poisson's ratio ν, so the constitutive equation takes the form

$$\begin{bmatrix} \sigma_{11} \\ \sigma_{22} \\ \sigma_{33} \\ \sigma_{12} \\ \sigma_{13} \\ \sigma_{23} \end{bmatrix} = \frac{E}{(1+\nu)(1-2\nu)} \begin{bmatrix} 1-\nu & \nu & \nu & 0 & 0 & 0 \\ \nu & 1-\nu & \nu & 0 & 0 & 0 \\ \nu & \nu & 1-\nu & 0 & 0 & 0 \\ 0 & 0 & 0 & 1-2\nu & 0 & 0 \\ 0 & 0 & 0 & 0 & 1-2\nu & 0 \\ 0 & 0 & 0 & 0 & 0 & 1-2\nu \end{bmatrix} \begin{bmatrix} \varepsilon_{11} \\ \varepsilon_{22} \\ \varepsilon_{33} \\ \varepsilon_{12} \\ \varepsilon_{13} \\ \varepsilon_{23} \end{bmatrix}. \tag{A1}$$

We avoid a numerical singularity by setting $\nu = 0.3$ and choose $E = 5$ kPa. Imposing (nearly) incompressible material constraints results in stresses that are very large and not physical. A total of 8000 strain vectors were

generated by randomly sampling each component of strain from a uniform distribution spanning $[-0.3, 0.3]$. Corresponding stress vectors were then computed for each strain vector using (A1). Furthermore, we initialize $\boldsymbol{S}_{\boldsymbol{x}}^{\varepsilon} = \mathbf{1}$ for all $\boldsymbol{x}$. All 8000 stress-strain pairs were used to train the MPN. The number of training pairs for the SN is fixed by the number of integration points in the FE mesh. Here, the FE mesh we created contained 14043 10-node tetrahedral elements, each with four integration points. Thus, there are 56172 total integration points in the phantom mesh and the same number of $\boldsymbol{x}$–$\boldsymbol{S}_{\boldsymbol{x}}^{\varepsilon}$ pairs.

References

Altahhan K N, Wang Y, Sobh N and Insana M F 2016 Indentation measurements to validate dynamic elasticity imaging methods *Ultrason. Imaging* **38** 332–45

Athanasiou A, Tardivon A, Tanter M, Sigal-Zafrani B, Bercoff J, Deffieux T, Gennisson J-L, Fink M and Neuenschwander S 2010 Breast lesions: quantitative elastography with supersonic shear imaging preliminary results *Radiology* **256** 297–303

Barbone P E and Gokhale N H 2004 Elastic modulus imaging: on the uniqueness and nonuniqueness of the elastography inverse problem in two dimensions *Inverse Problems* **20** 283

Barr R G 2010 Real-time ultrasound elasticity of the breast: initial clinical results *Ultrasound Q.* **26** 61–6

Barr R G, Destounis S, Lackey L B, Svensson W E, Balleyguier C and Smith C 2012 Evaluation of breast lesions using sonographic elasticity imaging: a multicenter trial *J. Ultrasound Med.* **31** 281–7

Bharat S, Fisher T G, Varghese T, Hall T J, Jiang J, Madsen E L, Zagzebski J A and Lee Jr F T 2008 Three-dimensional electrode displacement elastography using the siemens c7f2 foursight four-dimensional ultrasound transducer *Ultrasound Med. Biol.* **34** 1307–16

Bottenus N, Long W, Zhang H K, Jakovljevic M, Bradway D P, Boctor E M and Trahey G E 2016 Feasibility of swept synthetic aperture ultrasound imaging *IEEE Trans. Med. Imaging* **35** 1676–85

Deprez J-F, Brusseau E, Schmitt C, Cloutier G and Basset O 2009 3D estimation of soft biological tissue deformation from radio-frequency ultrasound volume acquisitions *Med. Image Anal.* **13** 116–27

Doyley M M 2012 Model-based elastography: a survey of approaches to the inverse elasticity problem *Phys. Med. Biol.* **57** R35

Fisher T G, Hall T J, Panda S, Richards M S, Barbone P E, Jiang J, Resnick J and Barnes S 2010 Volumetric elasticity imaging with a 2-D CMUT array *Ultrasound Med. Biol.* **36** 978–90

Fisher T G, Jiang J and Hall T 2007 Volumetric strain imaging *2007 Symp. Proc* pp 355–8 Oct 2007

Foroughi P, Kang H-J, Carnegie D A, van Vledder M G, Choti M A, Hager G D and Boctor E M 2013 A freehand ultrasound elastography system with tracking for *in vivo* applications *Ultrasound Med. Biol.* **39** 211–25

Ghaboussi J, Pecknold D A, Zhang M and Haj-ali R M 1998 Autoprogressive training of neural network constitutive models *Int. J. Numer. Methods Eng.* **42** 105–26

Ghaboussi J, Wu X S and Pecknold D A 2017 *Nonlinear Computational Solid Mechanics* (Boca Raton, FL: CRC Press) (https://doi.org/10.1201/9781315167329)

Gijsbertse K, Sprengers A M, Nillesen M M, Hansen H H, Lopata R, Verdonschot N and de Korte C L 2016 Three-dimensional ultrasound strain imaging of skeletal muscles *Phys. Med. Biol.* **62** 596

Hall T J, Bilgen M, Insana M F and Krouskop T A 1997 Phantom materials for elastography *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **44** 1355–65

Hashash Y M, Ghaboussi J, Fu Q and Marulanda C 2006 Constitutive soil behavior representation via artificial neural networks: a shift from soil models to soil behavior data *GeoCongress 2006: Geotechnical Engineering in the Information Technology Age* pp 1–6

Hashash Y M, Ghaboussi J and Jung S 2006 Characterizing granular material constitutive behavior using SelfSim with boundary load-displacement measurements *Earth and Space 2006: Engineering, Construction and Operations in Challenging Environment* pp 1–8

Hashemi H S and Rivaz H 2017 Global time-delay estimation in ultrasound elastography *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **64** 1625–36

Hendriks G A, Chen C, Hansen H H and de Korte C L 2018 3-D single breath-hold shear strain estimation for improved breast lesion detection and classification in automated volumetric ultrasound scanners *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **65** 1590–9

Hendriks G A, Holländer B, Menssen J, Milkowski A, Hansen H H and de Korte C L 2016 Automated 3D ultrasound elastography of the breast: a phantom validation study *Phys. Med. Biol.* **61** 2665

Hoerig C A data-driven approach to quasi-static ultrasonic elasticity imaging, Ph.D. dissertation University of Illinois, Urbana-Champaign (http://hdl.handle.net/2142/102440)

Hoerig C, Ghaboussi J and Insana M F 2017 An information-based machine learning approach to elasticity imaging *Biomech. Model. Mechanobiol.* **16** 805–22

Hoerig C, Ghaboussi J and Insana M F 2018 Cartesian neural network constitutive models for data-driven elasticity imaging *arXiv preprint arXiv:1809.04121* (https://arxiv.org/abs/1809.04121)

Hoerig C, Ghaboussi J and Insana M F 2019 Data-driven elasticity imaging using cartesian neural network constitutive models and the autoprogressive method *IEEE Trans. Med. Imaging* **38** 1150–60

Housden R J *et al* 2011 A new method for the acquisition of ultrasonic strain image volumes *Ultrasound Med. Biol.* **37** 434–41

Jung S and Ghaboussi J 2006 Characterizing rate-dependent material behaviors in self-learning simulation *Comput. Methods Appl. Mech. Eng.* **196** 608–19

Jung S and Ghaboussi J 2010 Inverse identification of creep of concrete from in situ load–displacement monitoring *Eng. Struct.* **32** 1437–45

Kim J, Ghaboussi J and Elnashai A S 2012 Hysteretic mechanical informational modeling of bolted steel frame connections *Eng. Struct.* **45** 1–11

Lee F-F, He Q and Luo J 2018 Electromagnetic tracking-based freehand 3D quasi-static elastography with 1D linear array: a phantom study *Phys. Med. Biol.* **63** 245006

Lindop J E, Treece G M, Gee A H and Prager R W 2006 3D elastography using freehand ultrasound *Ultrasound Med. Biol.* **32** 529–45

Liu B, Zheng Y, Huang G, Lin M, Shan Q, Lu Y, Tian W and Xie X 2016 Breast lesions: quantitative diagnosis using ultrasound shear wave elastography asystematic review and meta-analysis *Ultrasound Med. Biol.* **42** 835–47

Lubinski M A, Emelianov S Y and O'Donnell M 1999 Speckle tracking methods for ultrasonic elasticity imaging using short-time correlation *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **46** 82–96

Ophir J, Cespedes I, Ponnekanti H, Yazdi Y and Li X 1991 Elastography: a quantitative method for imaging the elasticity of biological tissues *Ultrason. Imaging* **13** 111–34

Papadacci C, Bunting E A and Konofagou E E 2016 3D quasi-static ultrasound elastography with plane wave in vivo *IEEE Trans. Med. Imaging* **36** 357–65

Papadacci C, Bunting E A, Wan E Y, Nauleau P and Konofagou E E 2016 3D myocardial elastography in vivo *IEEE Trans. Med. Imaging* **36** 618–27

Richards M S, Barbone P E and Oberai A A 2009 Quantitative three-dimensional elasticity imaging from quasi-static deformation: a phantom study *Phys. Med. Biol.* **54** 757

Riedmiller M and Braun H 1993 A direct adaptive method for faster backpropagation learning: the RPROP algorithm *IEEE Int. Conf. on Neural Networks* **1** 586–91

Sayed A, Layne G, Abraham J and Mukdadi O 2013 Nonlinear characterization of breast cancer using multi-compression 3D ultrasound elastography *in vivo Ultrasonics* **53** 979–91

Sayed A, Layne G, Abraham J and Mukdadi O M 2014 3-D visualization and non-linear tissue classification of breast tumors using ultrasound elastography *in vivo Ultrasound Med. Biol.* **40** 1490–502

Schaefer F *et al* 2011 Breast ultrasound elastography results of 193 breast lesions in a prospective study with histopathologic correlation *Eur. J. Radiol.* **77** 450–6

Szabo T L and Lewin P A 2013 Ultrasound transducer selection in clinical imaging practice *J. Ultrasound Med.* **32** 573–82

Treece G M, Lindop J E, Gee A H and Prager R W 2008 Freehand ultrasound elastography with a 3-D probe *Ultrasound Med. Biol.* **34** 463–74

Tyagi M, Goenezen S, Barbone P E and Oberai A A 2014 Algorithms for quantitative quasi-static elasticity imaging using force data *Int. J. Numer. Methods Biomed. Eng.* **30** 1421–36

Tyagi M, Wang Y, Hall T J, Barbone P E and Oberai A A 2017 Improving three-dimensional mechanical imaging of breast lesions with principal component analysis *Med. Phys.* **44** 4194–203

Wang Y, Nasief H G, Kohn S, Milkowski A, Clary T, Barnes S, Barbone P E and Hall T J 2017 Three-dimensional ultrasound elasticity imaging on an automated breast volume scanning system *Ultrason. Imaging* **39** 369–92

Zhi H, Ou B, Luo B-M, Feng X, Wen Y-L and Yang H-Y 2007 Comparison of ultrasound elastography, mammography and sonography in the diagnosis of solid breast lesions *J. Ultrasound Med.* **26** 807–15