A generic physics-informed neural network-based constitutive model for soft biological tissues

Minliang Liu[a], **Liang Liang**[b], **Wei Sun**[a,*]

[a]Tissue Mechanics Laboratory, The Wallace H. Coulter Department of Biomedical Engineering, Georgia Institute of Technology and Emory University, Atlanta, GA, United States of America

[b]Department of Computer Science, University of Miami, Coral Gables, FL, United States of America

Abstract

Constitutive modeling is a cornerstone for stress analysis of mechanical behaviors of biological soft tissues. Recently, it has been shown that machine learning (ML) techniques, trained by supervised learning, are powerful in building a direct linkage between input and output, which can be the strain and stress relation in constitutive modeling. In this study, we developed a novel generic physics-informed neural network material (NNMat) model which employs a hierarchical learning strategy by following the steps: (1) establishing constitutive laws to describe general characteristic behaviors of a class of materials; (2) determining constitutive parameters for an individual subject. A novel neural network structure was proposed which has two sets of parameters: (1) a class parameter set for characterizing the general elastic properties; and (2) a subject parameter set (three parameters) for describing individual material response. The trained NNMat model may be directly adopted for a different subject without re-training the class parameters, and only the subject parameters are considered as constitutive parameters. Skip connections are utilized in the neural network to facilitate hierarchical learning. A convexity constraint was imposed to the NNMat model to ensure that the constitutive model is physically relevant. The NNMat model was trained, cross-validated and tested using biaxial testing data of 63 ascending thoracic aortic aneurysm tissue samples, which was compared to expert-constructed models (Holzapfel-Gasser-Ogden, Gasser–Ogden–Holzapfel, and four-fiber families) using the same fitting and testing procedure. Our results demonstrated that the NNMat model has a significantly better performance in both fitting (R^2 value of 0.9632 vs 0.9019, p=0.0053) and testing (R^2 value of 0.9471 vs 0.8556, p=0.0203) than the Holzapfel–Gasser–Ogden model. The proposed NNMat model provides a convenient and general methodology for constitutive modeling.

*Correspondence to: The Wallace H. Coulter Department of Biomedical Engineering, Georgia Institute of Technology and Emory University, Technology Enterprise Park, Room 206 387 Technology Circle, Atlanta GA 30313-2412, United States of America. wei.sun@bme.gatech.edu (W. Sun).

Keywords

Constitutive modeling; Hyperelastic material; Machine learning; Neural network

1.   Introduction

Constitutive modeling is a cornerstone for stress analysis of mechanical behaviors of biological soft tissues [1–3]. Among the three key components required to solve a continuum biomechanics problem, i.e., the geometry (the domain of interest), the constitutive relations (how the material responds to applied loads under conditions of interest), and the applied loads (or associated boundary conditions), the identification of a robust constitutive model is probably the most challenging one to obtain and the key to success in this approach [4].

Currently, the approach to identify a robust constitutive model follows the DEICE procedure [4]: (1) **D**elineation of general characteristic behaviors, (2) **E**stablishment of an appropriate theoretical framework, (3) **I**dentification of specific functional forms of the constitutive relation, (4) **C**alculation of the values of the material parameters, and (5) **E**valuation of the predictive capability of the final constitutive relation. In this approach, a domain expert, (i.e., a biomechanicist with years of advanced training), plays a central role in the first 3 steps. A classic example is how Dr. Y. C. Fung discovered the famous Green-strain based, exponential form of the strain–energy function for soft tissues, iconized now as the Fung-elastic model [5,6]. Briefly, Fung showed that preconditioned soft tissue can be considered pseudo-elastic [5,6], and the slope of load–deflection curve is proportional to the load in uniaxial elongation tests of rabbit mesentery [7]. Consequently, an exponential function was used to account for the nonlinearity of the stress–strain curve for soft tissues. Indeed, the Green-strain based orthotropic form of the strain–energy function constructed by Fung provides excellent fitting capability with experimental data. To study biaxial mechanical properties of myocardial tissues, Humphrey et al. [8] performed constant invariant biaxial experiments, in which each of the strain invariant was independently varied, to infer specific functional forms of strain invariant-based constitutive equations. Based on the experimental observations, a polynomial form of the strain–energy function was devised [8]. To formulate a microstructurally-motivated constitutive model, Holzapfel et al. [9] modeled the arterial tissue as bi-layer fiber-reinforced composite, in which the contributions of a ground matrix and collagen fibers can be modeled separately in a strain–energy function.

Constitutive models [6,9–18] constructed by biomechanics experts have been widely adopted to model mechanical behaviors of soft tissues. By following the 3rd and 4th steps of the DEICE procedure, the specific formulations of these models usually contain several constitutive parameters that can be adjusted to describe constitutive behaviors of an individual subject (e.g. a tissue sample); therefore, the expert-constructed models can be used to describe constitutive behavior of a new subject (within the same class of materials) without deriving new constitutive equations. In addition, these expert-constructed constitutive models demonstrate excellent in-sample descriptive/fitting capability (e.g., R^2 value is high when fitting to mechanical testing data). However, their out-of-sample

predictive capability may be limited when new data (i.e., data that are not used in the fitting) is employed to assess their performance [19].

Recently, machine learning (ML) techniques, especially deep neural networks have led to revolutionary breakthrough in many applications [20–28], including recent works [29–34] in the field of biomechanics. Since ML techniques are capable of automatically discovering and capturing complex multi-dimensional input–output dependencies without the need of manually deriving specific functional forms, we hypothesize that a generic ML-based constitutive model can be developed and can have a similar, if not better, performance compared to the expert-constructed constitutive models.

Based on the universal function approximation theorems, a neural network with adequate capacities can approximate any continuous function with a small error [35–38]. Traditional feedforward fully-connected neural networks (FFNN) have been used to model the strain (input) and stress (output) relations [39–41]. However, such FFNN-based model uses all of its parameters (a.k.a. weights and biases) to construct the constitutive relation for an individual subject, which does not strictly follow the 3rd and 4th steps in the DEICE procedure; therefore, it often contains hundreds to thousands of constitutive parameters. Compared to an expert-constructed model, a FFNN-based constitutive model has three major disadvantages: (1) a large number of constitutive parameters with no physical meanings, in contrast to only a few constitutive parameters in an expert-constructed model. (2) An expert-constructed model can not only delineate and capture the general mechanical behaviors of a class of materials, but also can accurately model an individual subject (e.g. individual material responses) by fine tuning the constitutive parameters. A FFNN-based model, however, cannot capture general characteristic behaviors of a class of materials, i.e., it cannot utilize data from multiple subjects (e.g. tissue samples from many patients) for better modeling of an individual subject (e.g. a tissue sample from a single patient). The model parameters of FFNN-based models for different subjects are completely independent to each other. (3) a FFNN-based model cannot guarantee its convexity, which is important for ensuring the model is physically meaningful with unambiguous mechanical behaviors [9].

In this study, we developed a novel neural network-based material model (NNMat) which employs a physics constraint and a hierarchical learning strategy (Fig. 1): (1) establishing constitutive laws to describe general characteristic behaviors of a class of materials; (2) determining constitutive parameters for an individual subject. These two steps are equivalent to 3rd and 4th steps of the DEICE procedure. The neural network structure consists of two parameter sets corresponding to the two steps: (1) a "class" parameter set for characterizing the general elastic properties of the class of materials; and (2) a "subject" parameter set with three parameters for modeling individual material response. Skip connections are utilized in the neural network structure to facilitate hierarchical learning. Hence, the class parameters can function as the expert-constructed constitutive equations, and the NNMat model has only three constitutive parameters. The trained NNMat model may be directly adopted for a different subject without re-training the class parameters. The predictive capability of the proposed NNMat model is compared with the expert-constructed constitutive models (Holzapfel–Gasser–Ogden [9], Gasser–Ogden–Holzapfel [11], and four-fiber families [18]).

2. Constitutive modeling of soft biological tissues

Soft biological tissues comprise bundles of collagen fibers embedded in a ground matrix and can be regarded as fiber-reinforced composites. Constitutive modeling of the hyperelastic tissues is often achieved by specifying the strain energy density W as a function of deformation gradient W(**F**), where **F** represents the deformation gradient tensor. Microstructurally-motivated constitutive models have become increasingly utilized for soft tissues, in which the contributions of the matrix and collagen fibers can be modeled separately. In such models, the strain energy density function W is usually formulated based on strain invariants of the right Cauchy–Green tensor, $C = \mathbf{F}^T\mathbf{F}$. In this study, we consider a subclass of anisotropic responses, in which the strain energy density depends on four strain invariants: I_1, I_2, I_4 and I_6. The first two strain invariants I_1, I_2 are defined as

$$I_1 = tr(C), \ I_2 = \frac{1}{2}\left[I_1^2 - tr\left(C^2\right)\right] \tag{1}$$

For a fiber-reinforced composite material with two families of fibers, I_4 and I_6 are two additional pseudo-invariants that describes deformations in the preferred fiber directions

$$I_4 = \mathbf{a}_{01} \cdot (C\mathbf{a}_{01}), \ I_6 = \mathbf{a}_{02} \cdot (C\mathbf{a}_{02}) \tag{2}$$

where unit vectors $\mathbf{a}_{01}$ and $\mathbf{a}_{02}$ characterize two fiber directions in the reference configuration. Typically, these two fiber directions are assumed to be symmetric about an axis. $\mathbf{a}_{01} = (\cos \theta, \sin \theta, 0)$ and $\mathbf{a}_{02} = (\cos \theta, -\sin\theta, 0)$, where θ is the angle between the fiber direction and the axis of symmetry. In this study, the circumferential axis of the aorta was used as reference. Thus, I_4 and I_6 are equal to squares of the stretches in the fiber directions.

The stress–strain relation can be derived by differentiating the strain energy density W. For incompressible materials, the second Piola–Kirchhoff stress can be derived as [42]:

$$S = -pC^{-1} + 2W_1I + 2W_2(I_1I - C) + 2W_4\mathbf{a}_{01} \otimes \mathbf{a}_{01} + 2W_6\mathbf{a}_{02} \otimes \mathbf{a}_{02} \tag{3}$$

where I is the identity tensor, $W_k = \frac{\partial W}{\partial I_k}$, $k = 1, 2, 4, 6$ represent the derivatives of strain energy with respect to the strain invariants. p is the Lagrangian multiplier, which can be determined from boundary conditions. To characterize constitutive behavior, the relationship between strain invariants I_1, I_2, I_4, I_6 and W_1, W_2, W_4, W_6 needs to be established. Four nonlinear functions need to be constructed:

$$W_k = f_k(I_1, I_2, I_4, I_6), \ k = 1, 2, 4, 6 \tag{4}$$

2.1. Expert-constructed constitutive equations

Many expert-constructed models are available with specific formulations of W_k [9,11,13–17]. For comparison with our ML-based constitutive model, the Holzapfel–Gasser–Ogden(HGO) model [9] with two families of fibers was selected.

In the work by Holzapfel et al. [9], the total strain energy density function W can be additively split into isotropic W_{iso} and anisotropic $\mathbf{W}_{aniso}$ parts, according to

$$W(\mathbf{C}, \mathbf{a}_{01}, \mathbf{a}_{02}) = W_{iso}(\mathbf{C}) + \mathbf{W}_{aniso}(\mathbf{C}, \mathbf{a}_{01}, \mathbf{a}_{02}) \tag{5}$$

The isotropic matrix material is characterized by strain energy function of the neo-Hookean type

$$W_{iso}(\mathbf{C}) = C_{10}(I_1 - 3) \tag{6}$$

where C_{10} is a material parameter to describe the matrix response. To account for the strong stiffening effect of the collagen fiber recruitment, an exponential function is employed. The anisotropic contribution is given by

$$\mathbf{W}_{aniso}(\mathbf{C}, \mathbf{a}_{01}, \mathbf{a}_{02}) = \frac{k_1}{2k_2} \sum_{k=4,6} \left\{ exp\left[k_2(I_k - 1)^2 \right] - 1 \right\} \tag{7}$$

where k_1 is a positive material parameter that has the same unit of stress, while k_2 is a unitless material parameter. Hence, in the Holzapfel–Gasser–Ogden model [9], the relationship between strain invariants I_1, I_2, I_4, I_6 and strain energy derivatives W_1, W_2, W_4, W_6 can be obtained as:

$$\begin{aligned} W_1 &= C_{10} \\ W_2 &= 0 \\ W_4 &= k_1(I_4 - 1)exp\left[k_2(I_4 - 1)^2 \right] \\ W_6 &= k_1(I_6 - 1)exp\left[k_2(I_6 - 1)^2 \right] \end{aligned} \tag{8}$$

The four constitutive parameters $\{C_{10}, k_1, k_2, \theta\}$ can be determined through curve fitting to describe material properties of an individual subject. Stress–strain relation can be obtained using Eq. (3).

The Holzapfel–Gasser–Ogden model [9] has been extended to other forms. Using the generalized structural tensor (GST), Gasser et al. [11] constructed the following anisotropic contribution of the strain energy density function

$$\mathbf{W}_{aniso}(\mathbf{C}, \mathbf{a}_{01}, \mathbf{a}_{02}) = \frac{k_1}{2k_2} \sum_{k=4,6} \left[exp\left\{ k_2[\kappa I_1 + (1 - 3\kappa)I_k - 1]^2 \right\} - 1 \right] \tag{9}$$

where κ is a parameter describing dispersion of the fiber orientation. This model is known as the Gasser–Ogden–Holzapfel (GOH) model, which has five constitutive parameters $\{C_{10}, k_1, k_2, \kappa, \theta\}$. Hu et al. [18] proposed a four-fiber family model, which makes use of two additional invariants $I_\theta = \mathbf{a}_\theta \cdot (\mathbf{C}\mathbf{a}_\theta)$ and $I_z = \mathbf{a}_z \cdot (\mathbf{C}\mathbf{a}_z)$, along the circumferential ($\mathbf{a}_\theta = (1, 0, 0)$) and longitudinal ($\mathbf{a}_z = (0, 1, 0)$) directions, respectively. The anisotropic part of the strain energy density function is

$$W_{aniso}(\mathbf{C}, \mathbf{a}_{01}, \mathbf{a}_{02}) = \frac{k_1}{4k_2} \sum_{k=4,6} \left\{ exp\left[k_2(I_k - 1)^2 \right] - 1 \right\}$$
$$+ \frac{k_3}{4k_4} \sum_{l=\theta,z} \left\{ exp\left[k_4(I_l - 1)^2 \right] - 1 \right\}$$

(10)

Hence, the four-fiber model has six constitutive parameters $\{C_{10}, k_1, k_2, k_3, k_4, \theta\}$. In this study, the fitting and predictive capabilities of the Gasser–Ogden–Holzapfel model and the four-fiber family model are also demonstrated.

2.2. ML-based constitutive model

In this study, we developed a generic neural network-based material (NNMat) model (Fig. 2) with a novel neural network structure and a novel hierarchical learning strategy. The goal of the NNMat model is to establish the nonlinear mapping between I_1, I_2, I_4, I_6 and W_1, W_2, W_4, W_6 as described in Eq. (4) (W_2 may be non-zero). A physical constraint is added to the training process to ensure that convexity of the strain energy density is achieved by the NNMat model. Following the 3rd and 4th steps of the DEICE procedure, the NNMat model employs the hierarchical learning strategy: (1) constructing constitutive laws to describe general characteristic behaviors of a class of materials; (2) determining constitutive parameters for an individual subject. Therefore, the NNMat model has two sets of parameters: (1) a "class" parameter set for characterizing hyperelastic properties of the class of materials; and (2) a "subject" parameter set of three parameters for fitting mechanical response of an individual subject. NNMat models of different subjects will share the same class parameter set, but with different sets of subject parameters. In other words, each individual subject has an individual set of subject parameters and shares the same class parameter set with other subjects, assuming these subjects are from the same class of material.

Skip connections are increasingly utilized in deep learning research, e.g., additive skip connections in ResNet [43] and concatenative skip connections in DenseNet [44], to skip one or more layers in the neural network and connects the output of a previous layer to the next layers as the input. It can alleviate the vanishing-gradient problem by strengthening feature propagation and encouraging feature reuse [44]. In this study, a novel neural network structure with skip connections are proposed for the NNMat model to facilitate the hierarchical learning strategy.

The structure of the NNMat model is shown in Fig. 2. The class parameter set is processed by a fully-connected neural network, and two subject parameters are connected to each layer via concatenative skip connections, the subject parameter θ is incorporated in the unit directional vectors in Eq. (2). Therefore, the two subject parameters can have contributions to the output of each neuron in the hidden and output layers, which parametrizes the nonlinear mapping between I_1, I_2, I_4, I_6 and W_1, W_2, W_4, W_6. Comparing to only one subject parameter for one layer, skip connections may introduce more interactions between the subject parameters and the hidden/output layers. Specifically, the output of the ith neuron of the jth layer is a weighted sum of the input vector z^j, with weight w_i^j and bias b_i^j. In

addition, two subject parameters $\boldsymbol{m} = [m_1, m_2]^T$ are also connected (i.e. input) to each neuron with weight p_i^j utilizing concatenative skip connections,

$$u_i^j = \mathbf{w}_i^{j^T} z^j + \mathbf{p}_i^{j^T} \mathbf{m} + b_i^j \tag{11}$$

where the superscript j represents the layer index, and subscript i represents the neuron index. z^j represents the input to the neuron $\left[z_1^j, z_2^j, ..., z_{n_j}^j \right]^T$, and n_j denotes the number of neurons in the jth layer. $\left\{ \mathbf{W}_i^j, b_i^j, \mathbf{p}_i^j, i, j = 1, 2, ... \right\}$ is the class parameter set, which is contained in two hidden layers and the output layer, shown in blue color in Fig. 2. The linear combination u_i^j is nonlinearly transformed into the output z_i^{j+1} (the input to layer $j + 1$) using the softplus [45] activation function, given by

$$z_i^{j+1} = \log\left(1 + \exp\left(u_i^j \right) \right) \tag{12}$$

This function is a smooth version of the rectified linear unit (ReLU) [46]. The number of softplus units in the three layers are 128, 128 and 4, respectively. Consequently, in total, there are 18,188 class parameters $\left\{ \mathbf{W}_i^j, b_i^j, \mathbf{p}_i^j, i, j = 1, 2, ... \right\}$ in the NNMat model. The class parameters can function as the expert-constructed constitutive functional forms. Only the subject parameters $\{ m_1, m_2, \theta \}$ are considered as constitutive parameters.

After W_k are determined from the neural network, the second Piola–Kirchhoff stress can be computed using Eq. (3). The discrepancy between the experimental and predicted second Piola–Kirchhoff stress was measured by the mean squared error (MSE) loss function,

$$L_S = \sum_{m=1}^3 \frac{1}{N} \sum_{n=1}^N \left(S_m^{(n)} - \widehat{S}_m^{(n)} \right)^2 \tag{13}$$

where n is the data point index, N is the number of data points in the training dataset, $S_m^{(n)}$ and $\widehat{S}_m^{(n)}$ denote the kth predicted and experimental second Piola–Kirchhoff stress, respectively. m is the in-plane component index in the Voigt notation.

A physically-relevant strain energy field needs to be convex: the strain energy density should be at minimum (zero) when there is no deformation. In addition, since Newton's iterative method is often employed for solving nonlinear equations in commercial finite element (FE) packages, convexity of the strain energy density ensures that the material stiffness (Hessian) matrix will be positive definite and well-conditioned, which could stabilize the numerical solution [47]. For in-plane components, the Hessian matrix of the strain energy function can be expressed as:

$$\mathbf{H} = \begin{vmatrix} \dfrac{\partial S_{11}}{\partial E_{11}} & \dfrac{\partial S_{11}}{\partial E_{12}} & \dfrac{\partial S_{11}}{\partial E_{22}} \\[2ex] \dfrac{\partial S_{12}}{\partial E_{11}} & \dfrac{\partial S_{12}}{\partial E_{12}} & \dfrac{\partial S_{12}}{\partial E_{22}} \\[2ex] \dfrac{\partial S_{22}}{\partial E_{11}} & \dfrac{\partial S_{22}}{\partial E_{12}} & \dfrac{\partial S_{22}}{\partial E_{22}} \end{vmatrix} \tag{14}$$

For the strain energy function $W = \int S : dE$ to be path independent, it is required that $\mathbf{H}$ is symmetric. In the NNMat model, symmetry of $\mathbf{H}$ is enforced by a loss function at each data point n:

$$L_{c1}^{(n)} = \left| \frac{\partial S_{11}}{\partial E_{12}} - \frac{\partial S_{12}}{\partial E_{11}} \right| + \left| \frac{\partial S_{11}}{\partial E_{22}} - \frac{\partial S_{22}}{\partial E_{11}} \right| + \left| \frac{\partial S_{12}}{\partial E_{22}} - \frac{\partial S_{22}}{\partial E_{12}} \right| \tag{15}$$

Strict convexity requires positive definiteness of $\mathbf{H}$ to be satisfied for all possible strain values, which can be shown for expert-constructed model which has closed form solutions [48]. In the NNMat model, convexity of the strain energy density function is enforced by an additional loss function that ensures the positive semi-definiteness of Hessian matrix for all training stress–strain data points: for all $x \in \mathbb{R}^3$, $x^T \mathbf{H} x \geq 0$ needs to be satisfied. Using Sylvester's criterion of symmetric matrix, the requirement for positive semi-definiteness is that all of the principal minors must be non-negative [49,50]. Therefore, the positive semi-definiteness can be quantified by the following loss function for each data point n:

$$L_{c2}^{(n)} = \sum_{p=1}^{3} \sum_{q=1}^{\binom{3}{p}} \max\left(-\Delta_{p,q}, 0\right) \tag{16}$$

where $\Delta_{p,q}$ denote the qth principal minor of order p ($p = 1, 2, 3$) of the Hessian matrix. There are $\binom{3}{p}$ principal minors of order p. Hence, the loss function of the convexity constraints can be obtained by adding $L_{c1}^{(n)}$ and $L_{c2}^{(n)}$ with a weight α:

$$L_c = \frac{1}{N} \sum_{n=1}^{N} \left(L_{c1}^{(n)} + \alpha L_{c2}^{(n)} \right) \tag{17}$$

Since the loss functions are based on stress and its derivative, Eq. (3) needs to be included in the NNMat model for backpropagation, which is the last layer shown in yellow color in Fig. 2. The angle θ which defines the two fiber directions is another subject parameter in the NNMat. Hence, three subject parameters $\{m_1, m_2, \theta\}$ in the NNMat model can be adjusted for modeling properties of an individual subject. A bound constraint was imposed on the subject parameters to ensure that m_1, m_2 are in the range of −1 to 1 and θ is within −90° to 90°, which is realized by using the hyperbolic tangent function:

$$\widehat{m}_1 = \tanh(m_1)$$
$$\widehat{m}_2 = \tanh(m_2)$$
$$\widehat{\theta} = 90°\tanh(\theta)$$

(18)

where $\widehat{m}_1$, $\widehat{m}_2$ and $\widehat{\theta}$ represent the normalized parameters.

Therefore, the combined loss function for training the NNMat model is

$$L\left(\{m_1, m_2, \theta\}, \left\{\mathbf{W}_i^j, b_i^j, \mathbf{p}_i^j, i, j = 1, 2, \ldots\right\}\right) = L_S + \beta L_c$$

(19)

where β is another weight parameter. In this study, α and β were chosen using grid search in cross validation (Section 3.3). We note that the result may be refined using adaptive grid search. Other methods may be used for the hyperparameter optimization, e.g., random search [51] and Bayesian optimization [52]. Consequently, the combined loss function is a function of the (unknown) parameters in the NNMat model. The goal of training is to find the optimal values of the parameters in the NNMat model by minimizing the loss function on the training dataset.

We used a novel hierarchical training strategy to find the optimal parameters of the NNMat model. The class and subject parameters are determined in two sequential steps: (1) training and (2) fitting, which is equivalent to the 3rd and 4th steps of the DEICE procedure. During the model training, the class parameter set $\left\{\mathbf{W}_i^j, b_i^j, \mathbf{p}_i^j, i, j = 1, 2, \ldots\right\}$ is first optimized across all subjects in the training set, and the subject parameter set $\{m_1, m_2, \theta\}$ is then only optimized for the corresponding subject. During the model fitting, the class parameter set $\left\{\mathbf{W}_i^j, b_i^{\overline{j}}, \mathbf{p}_i^j, i, j = 1, 2, \ldots\right\}$ is fixed, i.e., it will no longer be updated through backpropagation; only the subject parameters $\{m_1, m_2, \theta\}$ are updated, which is similar to fitting an expert-constructed model. Therefore, the trained NNMat model may be directly adopted to a different subject without re-training the class parameters. When the trained NNMat model is used to characterize constitutive relation of a new subject, only the subject parameters need to be updated/fitted.

The NNMat model was implemented in PyTorch 1.0 [53]. Adam algorithm [54] was used for optimization to obtain the optimal parameters. Training, fitting and validation/testing of the NNMat model was run on a multi-GPU server (10-core CPU with 128 GB RAM, 4 × NVIDIA GeForce GTX 1080 Ti GPU).

3. Cross validation and testing

In this study, stress–strain data (63 patients) was split into two sets: a training and validation set (57 patients) and a testing set (6 patients). In the training mode, parameters in both the class set and subject set are updated. While in the fitting mode, the class parameters are fixed and only the three subject parameters are adjusted to optimal for an individual subject. For the testing/validation mode, all the parameters are fixed. The network structure and hyperparameters, e.g., α and β, were determined through cross validation using the training

and validation dataset. The performance of the NNMat model was evaluated using the additional testing dataset.

3.1. Planar biaxial testing data

In this study, we demonstrate the capability of the proposed NNMat model by using seven-protocol planar biaxial testing data [55] of ascending thoracic aortic aneurysm (ATAA) tissues from 63 patients/subjects that were published previously by our group [55,56]. Briefly, before planar biaxial testing, cryopreserved tissue samples were submerged in a 37 °C water bath until totally defrosted, following the two-stage slow thawing method to remove the cryopreservation agent [57]. The samples were trimmed into square-shaped specimens with a side length of 20~25 mm. Each specimen was subjected to biaxial tension with the circumferential (11) and longitudinal (22) directions aligned with the primary axes of the biaxial test fixture. A stress-controlled biaxial testing protocol was used. N denotes the nominal stress, and the ratio $N_{11} : N_{22}$ was kept constant. Each tissue specimen was preconditioned for at least 40 continuous cycles with $N_{11} : N_{22} = 1: 1$ to minimize hysteresis. Seven successive protocols were performed using ratios $N_{11}: N_{22} = 0.3: 1, 0.5: 1, 0.75: 1, 1: 1, 1: 0.75, 1: 0.5, 1: 0.3$. Fig. 3 shows representative biaxial testing results of two ATAA samples. Biaxial testing data was chosen since it contains hyperelastic properties under various in-plane stress ratios, which can be easily split into fitting and testing/validation dataset (i.e., data from six stress protocols for fitting the subject parameters and data from one protocol for testing/validation).

3.2. Fitting and testing of the expert-constructed models

Traditionally, for fitting an expert-constructed constitutive model to data, an error function is built based on difference between the experimental data and model predictions, and then constitutive parameters can be determined through nonlinear optimization. Typically, the model fitting process uses all experimental data (i.e. all seven stress protocols), and a coefficient of determination (R^2) as an accuracy metric is reported. However, since this R^2 metric corresponds to the in-sample prediction, the model's ability for out-of-sample prediction cannot be assessed, as shown in Schroeder et al. [19]. It is also reported [58] that a Fung-type hyperelastic model [59] with good in-sample fitting could result in erroneous out-of-sample stress predictions. As a consequence, the particular form of the constitutive model [59] needed to be modified by domain experts [58] to achieve a reasonable out-of-sample accuracy.

In this study, the predictive capabilities of the expert-constructed constitutive models in Section 2.1 were evaluated using cross validation. For each subject (i.e. a tissue sample from a patient), leave-one-out cross validation was performed: for each round, data from one stress protocol were selected for testing, and data from the remaining six stress protocols were used for fitting the model to obtain constitutive parameters. The process is repeated seven times for all protocols for each subject/patient in the testing dataset. Therefore, the averaged out-of-sample testing R^2 is used to evaluate the model performance, which provides a baseline to compare with that of the NNMat model. The fitting process was implemented in MATLAB using a nonlinear least square solver.

3.3. Cross validation and testing of the ML-based constitutive model

To assess performance of the NNMat model, cross validation was performed at two levels using the training and validation dataset of 57 patients. As can be seen in Fig. 4, the cross validation procedure consists of an outer loop and an inner loop which correspond to the patient level and stress protocol level, respectively. At the patient level, ten-fold cross validation was performed as follows: (1) split the patient data into ten groups, and each group contains the data from 5~6 patients; (2) in each round of the ten-fold cross validation, select one group for validation and use the remaining nine groups for training. The NNMat model parameters in both the class set and subject set are updated during the training stage. From the cross validation on a patient level, the performance of the NNMat model can be assessed for each individual patient in the training and validation set.

Similar to Section 3.2, leave-one-out cross validation was carried out for each patient on the stress protocol level: data from six protocols were used for fitting the subject parameters $\{m_1, m_2, \theta\}$, and then data from the remaining one protocol were used for assessing the model performance. The capability of the NNMat model to predict stress–strain response under various in-plane stress ratios can be evaluated. The weights α and β, and the number of training epochs were determined during cross validation.

After the hyperparameters were determined, the NNMat model was trained using the training and validation set (57 patients). Using the additional testing dataset of 6 patients, fitting and testing R^2 were evaluated for each patient using the same leave-one-out fashion on the stress protocol level (see Fig. 5). The averaged testing accuracy was used to evaluate performance of the NNMat model. The computing time for training the NNMat model is approximately 16~18 h using the training and validation set (73,614 stress–strain data points) with 10,000 epochs on a single GPU. Fitting of subject parameters for one patient can be completed in less than one minute. Using the trained and fitted NNMat model, stress computation can be achieved instantaneously.

4. Results

4.1. Cross validation

Using the training and validation set of 57 patients, grid search was performed to select the weights α and β. To reduce computational cost, the number of training epochs was set to be 1000 for the grid search. It is convenient to examine the convexity of a strain energy density function with respect to two in-plane components of the Green strain E_{11} and E_{22} while the shear component E_{12} is set to zero [47]. To examine its convexity, using the trained and fitted NNMat model, the second Piola–Kirchhoff stress S_{11} and S_{22} can be computed at a series of strain E_{11} and E_{22} values (–0.1 to 0.5). To evaluate whether the convexity condition is satisfied, using Sylvester's criterion [49,50], we define a convexity criterion $\varXi$ for each pair of E_{11} and E_{22}

$$\Xi(E_{11}, E_{22}) = \max\left(-\frac{\partial S_{11}}{\partial E_{11}}\frac{\partial S_{22}}{\partial E_{22}} + \frac{\partial S_{11}}{\partial E_{22}}\frac{\partial S_{22}}{\partial E_{11}}, 0\right) + \max\left(-\frac{\partial S_{11}}{\partial E_{11}}, 0\right)$$
$$+ \max\left(-\frac{\partial S_{22}}{\partial E_{22}}, 0\right) \tag{20}$$

For all strain values, the percentage of convexity criterion equal to zero ($\Xi = 0$) can be used to quantify convexity of the NNMat model. Therefore, we define a convexity index (CI) for a trained and fitted NNMat model

$$CI = \frac{1}{M}\sum_{E_{11}}\sum_{E_{22}}\mathbf{1}_c[\Xi(E_{11}, E_{22})] \tag{21}$$

where M represents the total number of E_{11} and E_{22} values. $\mathbf{1}_c (\Xi)$ is an indicator function, $\mathbf{1}_c (\Xi) = 1$ when $\Xi = 0$; otherwise, $\mathbf{1}_c (\Xi) = 0$. Hence, for all fitted NNMat models (using different patients and protocols in the validation set), the mean CI can be used to measure the convexity of the NNMat model. The results of grid search are reported in Table 1. The set of weights, $\alpha = 0.1$ and $\beta = 10^{-5}$, was selected because it yields good validation R^2 and an acceptable mean CI. The small α and β values may be explained by the fact that the values of lost terms are orders of magnitude different.

The number of training epochs was then treated as another hyperparameter. The NNMat model was cross validated using different number of epochs (1000, 5000, 10000, 15000), the results are listed in Table 2. The number of epochs 10000 was chosen because it resulted in the best performance.

4.2. Testing

The NNMat model was trained using the training and validation set of 57 patients. The expert-constructed models [9,11,18] (Section 2.1) and the trained NNMat model (Section 2.2) were fitted to biaxial data of the 6 ATAA patients in the testing set. Typical stress–strain results of the NNMat and the HGO model for a representative patient are shown in Fig. 6. For comparison, coefficients of determination (R^2) were computed for model predictions in terms of the Cauchy stress. For this particular patient (patient 59), it can be seen that the NNMat model is slightly more accurate than the HGO model. To compare the testing results of the NNMat and expert-constructed models, the fitting R^2 (using six protocols) and testing R^2 (using one protocol) of the four constitutive models for the testing patients are reported Table 3. The mean and standard deviation of fitting R^2 and testing R^2 are shown in Table 4. It is demonstrated that the NNMat model significantly outperforms the HGO model in fitting and predicting the ATAA biaxial data (p-value is 0.0053 for fitting and 0.0203 for testing). Power of the t-test is 0.9939 and 0.8329 for fitting and testing, respectively, which represents the probability that the null hypothesis is correctly rejected. The NNMat model has a slightly better performance than the GOH model which employs GST for fiber dispersion, but the result is non-significant. The NNMat model has a similar performance comparing to the four-fiber model which makes use of two additional invariants.

To visualize the convexity of strain energy density function, similar to Section 4.1, S_{11} and S_{22} were computed at a series of E_{11} and E_{22} values (–0.1 to 0.5). The strain energy density function can be calculated through trapezoidal numerical integration of $W = \int S : dE$. The resulting contours of strain energy function are plotted and examined for each patient. Strain energy functions of four representative patients are shown in Fig. 7, which are approximately convex (strict convexity was not proved). *CI* values for the fitted NNMat models can be evaluated using Eq. (21), the mean *CI* in the testing set is 100.00%.

4.3. Parametric study

To study the effect of the two subject parameters $\boldsymbol{m} = [m_1, m_2]^T$ on hyperelastic properties of the tissue, a trained NNMat model was employed (trained using the training and validation set) with $\theta = 0°$. The parameter space was sampled at various values of $\hat{m}_1$ and $\hat{m}_2$. with an interval of 0.01 in each dimension (from –1 to 1). Note that sampling was performed with the normalized parameters (i.e., $\hat{m}_1$ and $\hat{m}_2$). We evaluated tangent modulus $T = \frac{\Delta S_{ii}}{\Delta E_{ii}}$ under low strain ($E_{ii} = 0.1$) and high strain ($E_{ii} = 0.5$) in the circumferential (11) and longitudinal (22) directions. The results are shown in Fig. 8. It can be observed that the high modulus regions are concentrated in the upper right corner of the parameter space, where $\hat{m}_1$ and $\hat{m}_2$ are both at maximum. Thus, the larger $\hat{m}_1$ and $\hat{m}_2$ are, the higher stiffness the material is, which is physically meaningful.

5. Discussion

In this study, a novel generic physics-informed machine learning model was proposed for constitutive modeling of soft biological tissues. The proposed NNMat model utilizes a hierarchical learning strategy: it can learn from data of multiple subjects to improve its prediction for individuals. The structure of the NNMat model consists of a class parameter set for characterizing hyperelastic properties of a class of materials and a subject parameter set (three parameters) for fitting mechanical response of an individual subject. In the NNMat model, the subject parameter set is connected to the neural network via skip connections to facilitate hierarchical learning. In addition, a novel hierarchical training strategy was devised to determine the optimal parameters, which involve the determination of common mechanical properties for a class of material (training) and fitting of mechanical response of individual subjects (fitting). The modeling process is analogous to the DEICE procedure for expert-constructed models. Consequently, the NNMat model provides a convenient and general methodology for constitutive modeling of soft biological tissues. It inherits advantages of ML approaches: (1) construction of the NNMat model is an automatic process and does not involve any manual derivations, and (2) predictive capability of the NNMat model with 3 constitutive parameters is, superior to the 4-parameter HGO model; slightly better than the 5-parameter GOH model with consideration of fiber dispersion; and similar to the 6-parameter four-fiber model with two additional invariants. It also shares desired properties with the expert-constructed models: (1) only a few constitutive (subject) parameters $\{m_1, m_2, \theta\}$ are needed for modeling an individual sample response; (2) a physics constraint was enforced for convexity of the strain energy density function, which ensures that the constitutive relation is physically-relevant and numerically stable. Since the

NNMat model is differentiable, the elasticity tensor can be numerically computed, which makes the model suitable to be intergraded into a FE solver; (3) physical meanings of the constitutive (subject) parameters can be analyzed. Although training of the NNMat model takes approximately 16~18 h in our exemplary application. It should be noted that more time and efforts may be required for a domain expert to derive constitutive equations [8].

The current NNMat model was developed for modeling the hyperelastic properties of soft tissues and trained by using planar biaxial data of aortic wall tissues, it may also be extended to model history- or rate-dependent constitutive relations. For materials that exhibit history dependence, stress/strain history and internal state variables may be included as additional input variables. For rate-dependent material behavior, stress/strain rate may be incorporated similar to the traditional FFNN models [60,61].

A constitutive model with a large number of constitutive parameters often results in over-parameterization: different combinations of constitutive parameters are nonlinearly coupled which can lead to very similar mechanical response. This is undesirable from a data fitting perspective, in which the optimization problem is highly nonlinear, multivariate and non-convex, which can cause optimization difficulty known as the local optima [62]. In this case, the mechanical properties cannot be unambiguously represented by a set of constitutive parameters. Therefore, a constitutive model with fewer constitutive parameters is always preferred [63,64]. In contrast to the traditional FFNN models [39–41,60,61] that have hundreds to thousands of parameters, the proposed NNMat model with only three constitutive (subject) parameters is analogous to the expert-constructed models, which may be more suitable for practical applications. The NNMat model may also facilitate inverse estimation of constitutive parameters from *in vivo* clinical data, such as image-derived aortic wall motions [62,65–69], for which only the data within the physiological range can be obtained and therefore the identification solution may not be unique if there are too many constitutive parameters in the model.

Classically, the bias–variance trade-off, which implies that the model complexity must be limited to avoid overfitting, is often considered when training a ML model. In many applications, neural network size is significantly larger than data size, and training error approaches zero, which would be traditionally considered overfitted. Therefore, standard regularization procedures, e.g., weight decay and weight pruning, are often employed when training a neural network to prevent overfitting. However, surprisingly, the performance of the over-parameterized network on test set can be excellent [70]. This could be explained by a "double-descent" risk curve [70], which indicates that once the network complexity exceeds a threshold (i.e., being over-parameterized), the test risk starts to decrease (i.e. high performance on test set). The choice of whether to regularize the model can also depend on the available data size. In this study, the NNMat model has 18,188 class parameters and 3 subject parameters. In total, our biaxial data consists of 82,200 strain–stress data points, and there are 73,614 strain–stress data points in the training and validation dataset. In addition, it is shown recently that physics constraint may provide regularization effect in the physics-informed neural network [71], which may significantly reduce the required training data size [72]. Hence, our model may be "underfitting" the data, not complex enough. Nevertheless, the NNMat model has good fitting and testing performance on the testing data without using

standard regularization methods. Readers may choose to incorporate standard regularization procedures when implementing their own model, depending on the type of application and available data size.

It is often necessary to modify the general form of the constitutive relation for modeling of a new (structurally different) class of material, e.g. from modeling the arterial wall to the passive myocardium [10]. Although the exponential functional form of strain energy density has been widely adopted in many expert-constructed constitutive models of soft tissues [6,9–13] to describe the strain stiffening effect due to recruitment of collagen fibers following the work by Dr. Y.C. Fung [7], a number of studies [8,73,74] advocated polynomial strain energy functions. In general, the choice of constitutive model should be dependent on the type of tissues and loading ranges in specific applications, e.g., the exponential function without consideration of damage can lead to over-prediction of stress at large strain conditions [15,75–77]. Since the proposed NNMat model is generic (i.e., not specifically designed for ATAA tissues), we anticipated that its network structure can be applied to other tissue behaviors and characteristics, and its predictive capability is largely dependent upon the training data provided to it. With abundant training data to be collected (larger population, wider loading range) in the future, the predictive capability of the NNMat is expected to be improved.

6. Conclusions

In this study, a physics-informed machine learning model was proposed for constitutive modeling of soft biological tissues. A neural network material model (NNMat) with novel structure and hierarchical learning strategy is proposed. The NNMat model consists of two parameter sets: the class parameter set for characterizing the general elastic properties of a class of materials and the subject parameter set with three parameters for individual material response. Skip connections are utilized in the neural network structure to facilitate hierarchical learning. The proposed NNMat model can learn from the data of multiple subjects to improve its prediction for individuals. Physics constraints were enforced for convexity of the strain energy density function. From the tests we performed, both in-sample and out-of-sample accuracy metrics of the NNMat model are significantly higher than the expert-constructed model.

Acknowledgments

References

[1]. Taylor CA, Figueroa CA, Patient-specific modeling of Cardiovascular Mechanics, Annu. Rev. Biomed. Eng 11 (2009) 109–134. [PubMed: 19400706]

[2]. Chabanas M, Luboz V, Payan Y, Patient specific finite element model of the face soft tissues for computer-assisted maxillofacial surgery, Med. Image Anal 7 (2003) 131–151. [PubMed: 12868618]

[3]. Wittek A, Miller K, Kikinis R, Warfield SK, Patient-specific model of brain deformation: Application to medical image registration, J. Biomech 40 (2007) 919–929. [PubMed: 16678834]

[4]. Humphrey J, Introduction to Biomechanics, Springer.

[5]. Fung Y, Fronek K, Patitucci P, Pseudoelasticity of arteries and the choice of its mathematical expression, Amer. J. Physiol.-Heart Circulatory Physiol 237 (1979) H620–H631.

[6]. Fung Y.-c., Biomechanics: Mechanical Properties of Living Tissues, Springer Science & Business Media, 2013.

[7]. Fung Y, Elasticity of soft tissues in simple elongation, Amer. J. Physiol.-Leg. Content 213 (1967) 1532–1544.

[8]. Humphrey JD, Strumpf RK, Yin FCP, Determination of a constitutive relation for Passive Myocardium: I. A new Functional form, J. Biomech. Eng 112 (1990) 333–339. [PubMed: 2214717]

[9]. Holzapfel GA, Gasser TC, Ogden RW, A new constitutive Framework for Arterial Wall mechanics and a comparative study of material models, J. Elast. Phys. Sci. Solids 61 (2000) 1–48.

[10]. Holzapfel GA, Ogden RW, Constitutive modelling of passive myocardium: a structurally based framework for material characterization, Phil. Trans. R. Soc. A 367 (2009) 3445–3475. [PubMed: 19657007]

[11]. Gasser TC, Ogden RW, Holzapfel GA, Hyperelastic modelling of arterial layers with distributed collagen fibre orientations, J. R. Soc. Interface 3 (2006) 15–35. [PubMed: 16849214]

[12]. Demiray H, A note on the elasticity of soft biological tissues, J. Biomech. 5 (1972) 309–311. [PubMed: 4666535]

[13]. Baek S, Gleason RL, Rajagopal KR, Humphrey JD, Theory of small on large: Potential utility in computations of fluid–solid interactions in arteries, Comput. Methods Appl. Mech. Engrg. 196 (2007) 3070–3078.

[14]. Li K, Ogden RW, Holzapfel GA, An exponential constitutive model excluding fibres under compression: Application to extension–inflation of a residually stressed carotid artery, Math. Mech. Solids 23 (2018) 1206–1224.

[15]. Li K, Holzapfel GA, Multiscale modeling of fiber recruitment and damage with a discrete fiber dispersion method, J. Mech. Phys. Solids 126 (2019) 226–244.

[16]. Li K, Ogden RW, Holzapfel GA, A discrete fibre dispersion method for excluding fibres under compression in the modelling of fibrous tissues, J. R. Soc. Interface 15 (2018) 20170766. [PubMed: 29386399]

[17]. Li K, Ogden RW, Holzapfel GA, Modeling fibrous biological tissues with a general invariant that excludes compressed fibers, J. Mech. Phys. Solids 110 (2018) 38–53.

[18]. Hu JJ, Baek S, Humphrey JD, Stress–strain behavior of the passive basilar artery in normotension and hypertension, J. Biomech. 40 (2007) 2559–2563. [PubMed: 17207488]

[19]. Schroeder F, Polzer S, Slažanský M, Man V, Skácel P, Predictive capabilities of various constitutive models for arterial tissue, J. Mech. Behav. Biomed. Mater 78 (2018) 369–380. [PubMed: 29220821]

[20]. He K, Zhang X, Ren S, Sun J, Delving deep into Rectifiers: Surpassing Human-Level performance on ImageNet classification, in: IEEE International Conference on Computer Vision, 2015.

[21]. Kokkinos I, Pushing the Boundaries of Boundary detection using Deep Learning, in: Int.l Conf. on Learning Representations, 2016.

[22]. Taigman Y, Yang M, Ranzato MA, Wolf L, DeepFace: Closing the Gap to Human-Level performance in Face verification, in: IEEE Conference on Computer Vision and Pattern Recognition, 2014.

[23]. He K, Zhang X, Ren S, Sun J, Deep Residual learning for Image Recognition, in: IEEE Conference on Computer Vision and Pattern Recognition, 2016.

[24]. Krizhevsky A, Sutskever I, Hinton GE, ImageNet classification with Deep convolutional Neural Networks, Neural Inf. Process. Syst (2012).

[25]. LeCun Y, Bengio Y, Hinton GE, Deep learning, Nature 521 (2015) 436–444. [PubMed: 26017442]

[26]. Wu Y, Schuster M, Chen Z, et al., Google's neural machine translation system: Bridging the Gap between Human and machine translation, Comput. Res. Repos (2016) abs/1609.08144.

[27]. Hannun A, Case C, Casper J, et al., Deep Speech: Scaling up end-to-end speech recognition, Comput. Res. Repos (2014) abs/1412.5567.

[28]. He X, Avril S, Lu J, Machine learning prediction of Tissue strength and local Rupture risk in Ascending thoracic Aortic Aneurysms, Mol. Cell. Biomech 16 (2019) 50–52.

[29]. Dabiri Y, Van der Velden A, Sack KL, Choy JS, Kassab GS, Guccione JM, Prediction of Left Ventricular mechanics using Machine learning, Front. Phys 7 (2019).

[30]. Romaszko L, Borowska A, Lazarus A, Gao H, Luo X, Husmeier D, in: Ladde G, Jeske D (Eds.), Direct Learning Left Ventricular Meshes from CMR Images, 2019, p. 25.

[31]. Luo Y, Fan Z, Baek S, Lu J, Machine learning–aided exploration of relationship between strength and elastic properties in ascending thoracic aneurysm, Int. J. Numer. Methods Biomed. Eng 34 (2018) e2977.

[32]. Do HN, Ijaz A, Gharahi H, Zambrano B, Choi J, Lee W, Baek S, Prediction of Abdominal Aortic Aneurysm Growth using dynamical Gaussian process Implicit surface, IEEE Trans. Biomed. Eng 66 (2019) 609–622. [PubMed: 29993480]

[33]. Jiang Z, Do HN, Choi J, Lee W, Baek S, A deep learning approach to Predict Abdominal Aortic Aneurysm expansion using longitudinal Data, Front. Phys 7 (2020).

[34]. Cilla M, Pérez-Rey I, Martínez MA, Peña E, Martínez J, On the use of machine learning techniques for the mechanical characterization of soft biological tissues, Int. J. Numer. Methods Biomed. Eng 34 (2018) e3121.

[35]. Csáji BC, Approximation with artificial neural networks, in: Faculty of Sciences, Vol. 24, Etvs Lornd University, Hungary, 2001, p. 48.

[36]. Leshno M, Lin VY, Pinkus A, Schocken S, Multilayer feedforward networks with a nonpolynomial activation function can approximate any function, Neural Netw 6 (1993) 861–867.

[37]. Cybenko G, Approximation by superpositions of a sigmoidal function, Math. Control Signals Systems 2 (1989) 303–314.

[38]. Hornik K, Approximation capabilities of multilayer feedforward networks, Neural Netw 4 (1991) 251–257.

[39]. Yagawa G, Okuda H, Neural networks in computational mechanics, Arch. Comput. Methods Eng 3 (1996) 435.

[40]. Ghaboussi J, Pecknold DA, Zhang M, Haj-Ali RM, Autoprogressive training of neural network constitutive models, Internat. J. Numer. Methods Engrg 42 (1998) 105–126.

[41]. Ghaboussi J, Sidarta DE, New nested adaptive neural networks (NANN) for constitutive modeling, Comput. Geotech 22 (1998) 29–52.

[42]. Holzapfel GA, Ogden RW, On planar biaxial tests for anisotropic nonlinearly elastic solids. A continuum mechanical framework, Math. Mech. Solids 14 (2009) 474–489.

[43]. He K, Zhang X, Ren S, Sun J, Deep residual learning for image recognition, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016, pp. 770–778.

[44]. Huang G, Liu Z, Van Der Maaten L, Weinberger KQ, Densely connected convolutional networks, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017, pp. 4700–4708.

[45]. Glorot X, Bordes A, Bengio Y, Deep sparse rectifier neural networks, in: Proceedings of the Fourteenth International Conference on Artificial Intelligence and Statistics, 2011, pp. 315–323.

[46]. Nair V, Hinton GE, Rectified linear units improve restricted boltzmann machines, in: Proceedings of the 27th International Conference on Machine Learning (ICML-10), 2010, pp. 807–814.

[47]. Sun W, Sacks MS, Finite element implementation of a generalized Fung-elastic constitutive model for planar soft tissues, Biomech. Model. Mechanobiol 4 (2005) 190–199. [PubMed: 16075264]

[48]. Balzani D, Neff P, Schröder J, Holzapfel GA, A polyconvex framework for soft biological tissues. Adjustment to experimental data, Int. J. Solids Struct 43 (2006) 6052–6070.

[49]. Meyer CD, Matrix Analysis and Applied Linear Algebra, Siam, 2000.

[50]. Prussing JE, The principal minor test for semidefinite matrices, J. Guid. Control Dyn 9 (1986) 121–122.

[51]. Bergstra J, Bengio Y, Random search for hyper-parameter optimization, J. Mach. Learn. Res 13 (2012) 281–305.

[52]. Snoek J, Larochelle H, Adams RP, Practical bayesian optimization of machine learning algorithms, in: Advances in Neural Information Processing Systems, 2012, pp. 2951–2959.

[53]. Paszke A, Gross S, Chintala S, Chanan G, Yang E, DeVito Z, Lin Z, Desmaison A, Antiga L, Lerer A, Automatic differentiation in pytorch, 2017.

[54]. Kingma DP, Ba J, Adam: A method for stochastic optimization, 2014, arXiv preprint arXiv:1412.6980.

[55]. Pham T, Martin C, Elefteriades J, Sun W, Biomechanical characterization of ascending aortic aneurysm with concomitant bicuspid aortic valve and bovine aortic arch, Acta Biomater. 9 (2013) 7927–7936. [PubMed: 23643809]

[56]. Martin C, Sun W, Pham T, Elefteriades J, Predictive biomechanical analysis of ascending aortic aneurysm rupture potential, Acta Biomater 9 (2013) 9392–9400. [PubMed: 23948500]

[57]. Bia D, Pessana F, Armentano R, Pérez H, Graf S, Zócalo Y, Saldías M, Perez N, Alvarez O, Silva W, Machin D, Sueta P, Ferrin S, Acosta M, Alvarez I, Cryopreservation procedure does not modify human carotid homografts mechanical properties: an isobaric and dynamic analysis, Cell Tissue Bank. 7 (2006) 183–194. [PubMed: 16933040]

[58]. Sun W, Sacks MS, Sellaro TL, Slaughter WS, Scott MJ, Biaxial mechanical response of Bioprosthetic Heart Valve Biomaterials to high In-plane Shear, J. Biomech. Eng 125 (2003) 372–380. [PubMed: 12929242]

[59]. Sacks MS, A method for planar Biaxial mechanical testing that Includes In-Plane Shear, J. Biomech. Eng 121 (1999) 551–555. [PubMed: 10529924]

[60]. Jung S, Ghaboussi J, Neural network constitutive model for rate-dependent materials, Comput. Struct 84 (2006) 955–963.

[61]. Stoffel M, Bamer F, Markert B, Neural network based constitutive modeling of nonlinear viscoplastic structural response, Mech. Res. Commun 95 (2019) 85–88.

[62]. Wittek A, Karatolios K, Bihari P, Schmitz-Rixen T, Moosdorf R, Vogt S, Blase C, In vivo determination of elastic properties of the human aorta based on 4D ultrasound data, J. Mech. Behav. Biomed. Mater 27 (2013) 167–183. [PubMed: 23668998]

[63]. Zeinali-Davarani S, Choi J, Baek S, On parameter estimation for biaxial mechanical behavior of arteries, J. Biomech 42 (2009) 524–530. [PubMed: 19159887]

[64]. Guan D, Ahmad F, Theobald P, Soe S, Luo X, Gao H, On the AIC-based model reduction for the general Holzapfel–Ogden myocardial constitutive law, Biomech. Model. Mechanobiol 18 (2019) 1213–1232. [PubMed: 30945052]

[65]. Wittek A, Karatolios K, Fritzen C-P, Bereiter-Hahn J, Schieffer B, Moosdorf R, Vogt S, Blase C, Cyclic three-dimensional wall motion of the human ascending and abdominal aorta characterized by time-resolved three-dimensional ultrasound speckle tracking, Biomech. Model. Mechanobiol 15 (2016) 1375–1388. [PubMed: 26897533]

[66]. Liu M, Liang L, Sun W, Estimation of in vivo mechanical properties of the aortic wall: A multi-resolution direct search approach, J. Mech. Behav. Biomed. Mater 77 (2017) 649–659. [PubMed: 29101897]

[67]. Liu M, Liang L, Sulejmani F, Lou X, Iannucci G, Chen E, Leshnower B, Sun W, Identification of in vivo nonlinear anisotropic mechanical properties of ascending thoracic aortic aneurysm from patient-specific CT scans, Sci. Rep 9 (2019) 12983. [PubMed: 31506507]

[68]. Liu M, Liang L, Sun W, A new inverse method for estimation of in vivo mechanical properties of the aortic wall, J. Mech. Behav. Biomed. Mater 72 (2017) 148–158. [PubMed: 28494272]

[69]. Liu M, Liang L, Sun W, Estimation of in vivo constitutive parameters of the aortic wall using a machine learning approach, Comput. Methods Appl. Mech. Engrg 347 (2019) 201–217.

[70]. Belkin M, Hsu D, Ma S, Mandal S, Reconciling modern machine-learning practice and the classical bias–variance trade-off, Proc. Natl. Acad. Sci 116 (2019) 15849–15854. [PubMed: 31341078]

[71]. Raissi M, Perdikaris P, Karniadakis GE, Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations, J. Comput. Phys 378 (2019) 686–707.

[72]. Liu D, Wang Y, A Dual-Dimer method for Training physics-constrained Neural Networks with Minimax architecture, 2020, arXiv preprint arXiv:2005.00615.

[73]. Vaishnav RN, Young JT, Janicki JS, Patel DJ, Nonlinear Anisotropic Elastic properties of the Canine Aorta, Biophys. J 12 (1972) 1008–1027. [PubMed: 5044576]

[74]. Raghavan ML, Vorp DA, Toward a biomechanical tool to evaluate rupture potential of abdominal aortic aneurysm: identification of a finite strain constitutive model and evaluation of its applicability, J. Biomech 33 (2000) 475–482. [PubMed: 10768396]

[75]. Weisbecker H, Pierce DM, Regitnig P, Holzapfel GA, Layer-specific damage experiments and modeling of human thoracic and abdominal aortas with non-atherosclerotic intimal thickening, J. Mech. Behav. Biomed. Mater 12 (2012) 93–106. [PubMed: 22659370]

[76]. Pierce DM, Maier F, Weisbecker H, Viertler C, Verbrugghe P, Famaey N, Fourneau I, Herijgers P, Holzapfel GA, Human thoracic and abdominal aortic aneurysmal tissues: Damage experiments, statistical analysis and constitutive modeling, J. Mech. Behav. Biomed. Mater 41 (2015) 92–107. [PubMed: 25460406]

[77]. Schmidt T, Balzani D, Holzapfel GA, Statistical approach for a continuum description of damage evolution in soft collagenous tissues, Comput. Methods Appl. Mech. Engrg 278 (2014) 41–61.
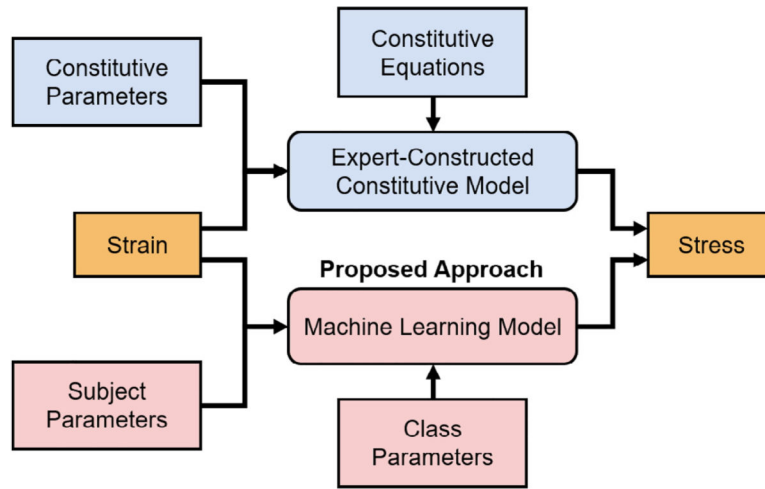
Fig. 1.
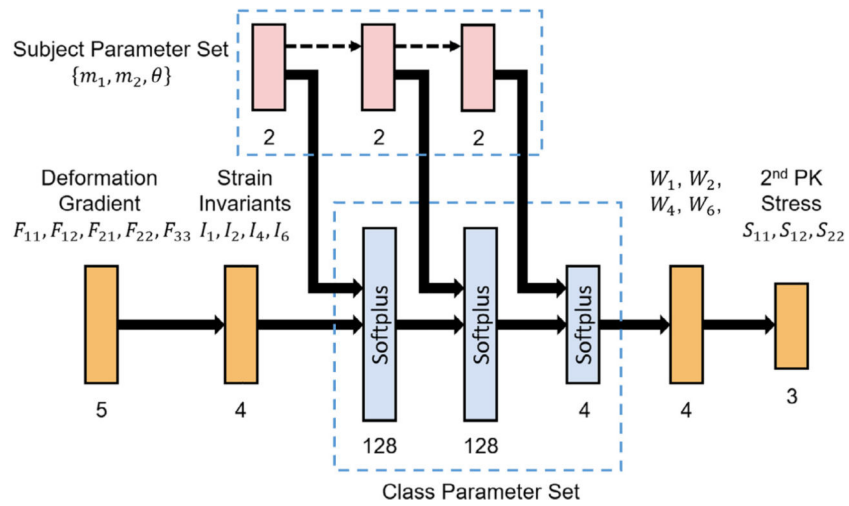The proposed machine learning-based constitutive model.

Fig. 2.
The novel neural network material model (NNMat) with a physics constraint and two parameter-set structures: the class parameter set (blue) and the subject parameter set (red), dashed arrows indicate skip connections. The subject parameter set consists of three constitutive parameters: $\{m_1, m_2, \theta\}$.
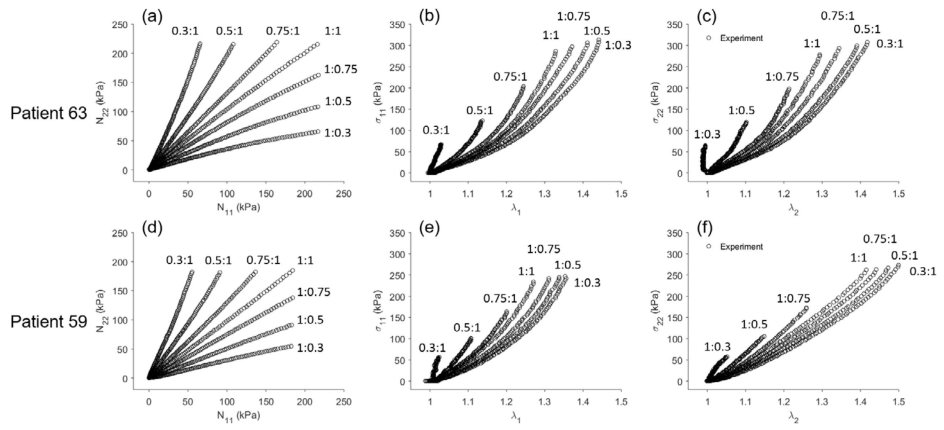
Fig. 3.
Representative stress–stretch results of the seven-protocol planar biaxial testing of two ATAA tissues. Each row represents one patient. (a) and (d): seven successive nominal stress ratios. (b) and (e): circumferential stress–stretch data. (c) and (f): longitudinal stress–stretch data.
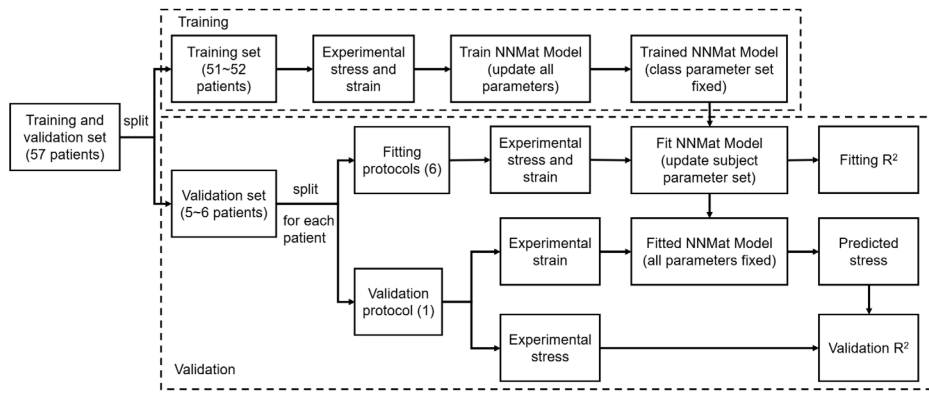
Fig. 4.

Cross validation of the NNMat model on patient level and stress protocol level. Ten-fold cross validation was performed for different patient groups. For each patient, leave-one-out cross validation was performed with different stress protocols.
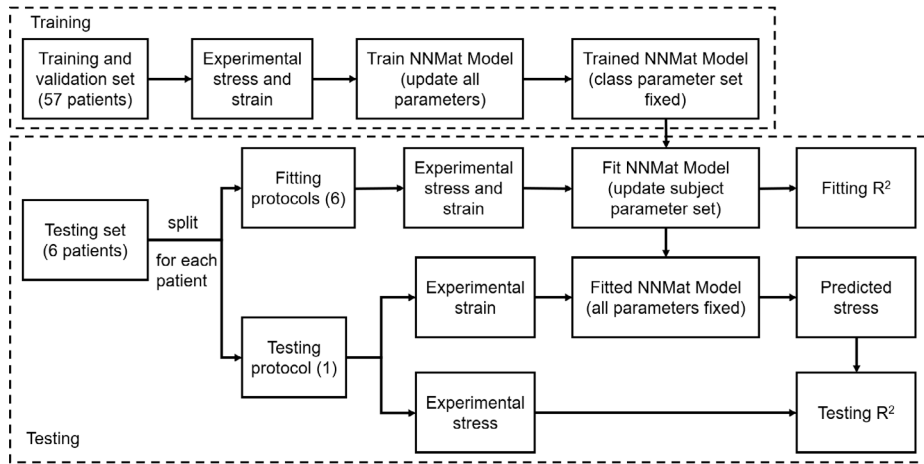
Fig. 5.
Evaluating accuracy of the NNMat model using an additional testing dataset of 6 patients.
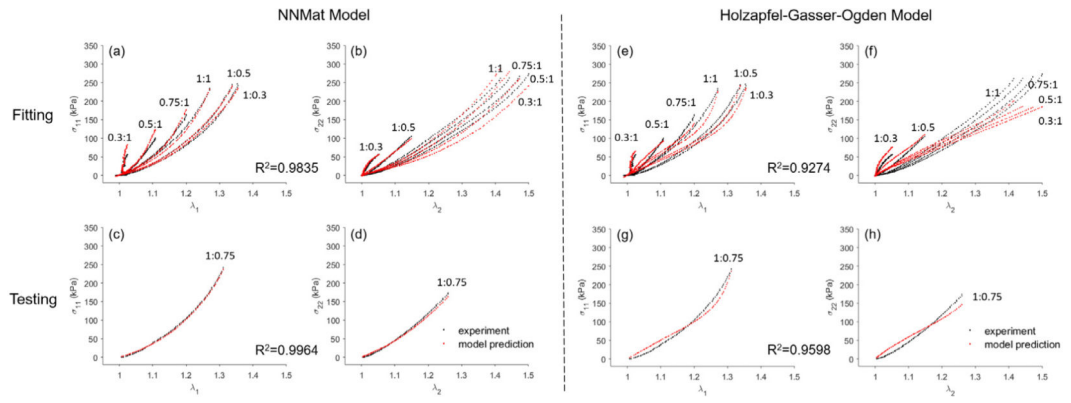
Fig. 6.
Representative Cauchy stress stretch results of the NNMat model and the Holzapfel–Gasser–
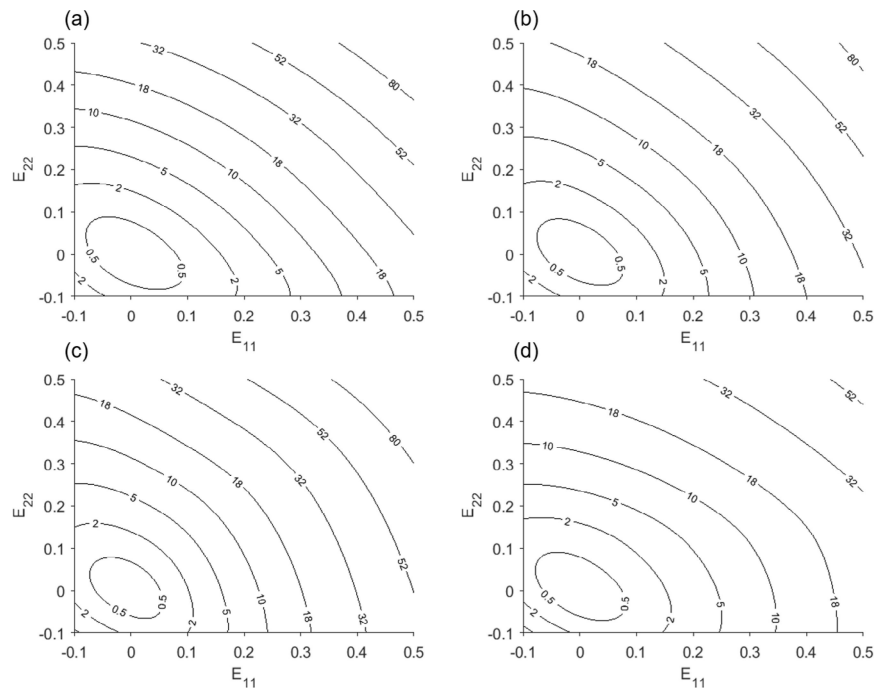Ogden model [9] (Section 2.1) in fitting and testing for patient 59.

Fig. 7.
Strain energy density function with respect to E_{11} and E_{22} with $E_{12} = 0$. Four representative patients are plotted. (a): patient 58, (b): patient 59, (c): patient 60 and (d): patient 62.
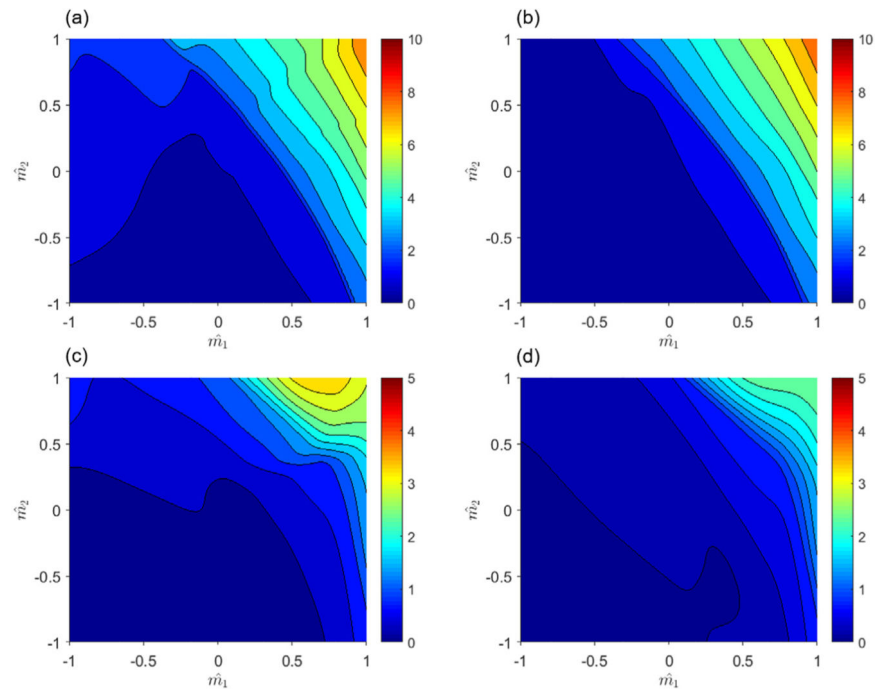
Fig. 8.
Contour plots of tangent moduli in the subject parameter space ($\theta = 0°$, trained using patients 7~63). (a) and (b): circumferential; (c) and (d): longitudinal. (a) and (c): tangent modulus at high strain region (0.5); (d) and (d): tangent modulus at low strain region (0.1). Units of tangent moduli are in MPa.

Table 1

Fitting and validation results obtained from grid search with different α and β values. The number of training epochs was set to be 1000.

(β, α)	Mean fitting R^2	Mean validation R^2	Mean CI
$(10^{-3}, 1)$	0.8280	0.7884	99.43%
$(10^{-3}, 0.1)$	0.8628	0.8310	94.56%
$(10^{-3}, 0.01)$	0.8603	0.8204	95.40%
$(10^{-5}, 1)$	0.8668	0.8397	95.93%
$(10^{-5}, 0.1)$	0.8765	0.8586	96.08%
$(10^{-5}, 0.01)$	0.8789	0.8564	95.07%
$(10^{-7}, 1)$	0.8833	0.8620	95.25%
$(10^{-7}, 0.1)$	0.8865	0.8645	93.84%
$(10^{-7}, 0.01)$	0.8872	0.8641	92.90%

Table 2

Fitting and validation results obtained using different number of training epochs.

Number of epochs	Mean fitting R^2	Mean validation R^2	Mean CI
1000	0.8765	0.8586	96.08%
5000	0.8998	0.8691	97.40%
10 000	0.9359	0.9139	99.10%
15 000	0.9192	0.8940	96.24%

Table 3

Fitting and testing R^2 predicted by the NNMat model and expert-constructed models using the testing dataset.

NNMat (3 parameters)			Holzapfel–Gasser–Ogden (4 parameters)		
Patient ID	Fitting R^2	Testing R^2	Patient ID	Fitting R^2	Testing R^2
58	0.9832	0.9747	58	0.9006	0.8737
59	0.9755	0.9749	59	0.9303	0.9197
60	0.9736	0.9511	60	0.9247	0.9171
61	0.9672	0.9658	61	0.8903	0.8094
62	0.8902	0.8272	62	0.8475	0.7297
63	0.9895	0.9891	63	0.9178	0.8838

Gasser–Ogden–Holzapfel (5 parameters)			Four-fiber families (6 parameters)		
Patient ID	Fitting R^2	Testing R^2	Patient ID	Fitting R^2	Testing R^2
58	0.9683	0.9523	58	0.9710	0.9620
59	0.9638	0.9491	59	0.9821	0.9784
60	0.9395	0.9187	60	0.9779	0.9605
61	0.9422	0.9185	61	0.9767	0.9661
62	0.8780	0.8103	62	0.8881	0.7803
63	0.9748	0.9627	63	0.9764	0.9705

Table 4

Fitting and testing R^2 (mean ± standard deviation) obtained by the NNMat model and the expert-constructed models. The p-values are computed using Behrens–Fisher two-sample t-test with the null hypothesis that R^2 of the expert-constructed model is greater than that of the NNMat model. The results indicate that performance of the NNMat model is much better than the HGO model.

Model	Fitting R^2	*p*-value	Testing R^2	*p*-value
NNMat	0.9632 ± 0.0366		0.9471 ± 0.0601	
Holzapfel–Gasser–Ogden	0.9019 ± 0.0306	0.0053	0.8556 ± 0.0735	0.0203
Gasser–Ogden–Holzapfel	0.9445 ± 0.0355	0.1944	0.9186 ± 0.0561	0.2076
Four-fiber families	0.9620 ± 0.0364	0.4782	0.9363 ± 0.0767	0.3956