

Pipe aspiration method for characterizing multi-layer soft materials based on statistical learning framework

zhenghui guo (zg27) Junyuan Sheng (js213)

ABSTRACT Characterizing the mechanical properties of soft biomaterials is of great significance in interpretation of biological mechanisms, disease diagnosis, and tissue engineering applications. A fairly large number of biomaterials have multi-layered structures, which involve numerous geometric and physical parameters, leading the problems complicated. In our research, we propose a pipette aspiration method for characterizing the mechanical properties of film-substrate soft materials under statistical learning framework. Starting from dimensional analysis, we extract dimensionless parameters of this problem. Combined with FEM simulation, the large-scale data set is gathered. Machine learning methods (deep neural network and XGBoost) are carried out to build high predictive models. Our model results exhibit good accuracy and generalization ability.

1 Introduction

The pipette aspiration method has been widely used in characterizing the local mechanical properties of soft materials [1-4]. During tests, the soft materials are aspirated into a glass pipette with a small diameter, and the aspiration length is determined by applying pressure to it. The material properties can be extracted from the pressure–aspiration length curves by different approaches such as analytical models, numerically derived methods, inverting the material properties with finite-element approaches, or inverting by measuring the wavelength in unstable configurations [5-10]. When geometric and material nonlinearities are present in the pipette aspiration test, it is difficult to develop an analytical solution correlating the experimental responses with material properties. Therefore, inverse analyses used to interpret experimental data are usually based on finite element simulations [11]. A micropipette aspiration method combined with finite element method was used to investigate the mechanical properties of multilayer soft biological materials [12].

Another challenging problem is that the film-substrate soft materials have complex structures and involve parameters coupling, leading lack of effective models to describe the mechanical behaviors. Data-driven methods with machine learning algorithms have been demonstrated to be a powerful tool and applied in various relative fields [13-18]. The deep neural network (DNN) method has been proven to be efficient in discovering meaningful structures in data. This method can extract high-dimensional features from original data by learning how to transform the high-dimensional feature vector into a specific target by using a sufficient amount of training data [19]. The tree-based boosting method is also widely used and highly effective. XGBoost is regarded as a scalable, end-to-end tree boosting system that is used by many data scientists to tackle a wide range of machine learning problems [20].

In this work, we propose machine learning to solve this problem. For the classic film-substrate system, we obtain the data set with large-scale finite element calculations. Dimensional analysis is employed to extract the features of the problem. To quantify the mechanical properties of the film, machine learning methods are applied to reveal the relationship between the physical quantities in the pipette problem. We employ Deep Neural Network (DNN) and XGBoost to train machine learning models and establish the relationship between each physical quantity.

2 Method

2.1 Dimensional analysis on film-substrate soft material

During the pipette aspiration experiment, we put a pipette attach to the top of material, and then exert negative pressure on the soft material, shown in Figure 1. The mechanical properties of the tested material can be extracted based on the data of the pressure and aspiration length. In this research, we focus on bilayer soft material (film-substrate system). R and t_p are the inner radius and wall thickness of the pipette. P is the negative pressure applied in the pipette, and t is the thickness of top-layer film. The aspiration length is l . μ_f and μ_s are the initial shear modulus of the film and the substrate, respectively. Since the inner radius of the pipette is much smaller than the tested material size, the material can be regarded as semi-infinite. The following finite element validation shows that when $t_p/R > 0.3$, the influence of pipe wall thickness t_p can be ignored.

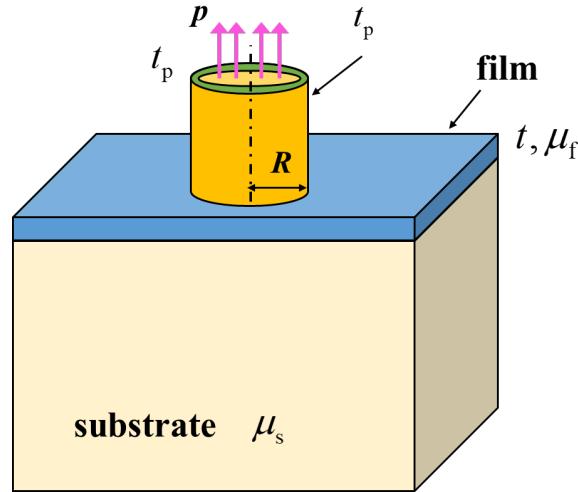


Figure. 1 Schematic of pipette aspiration test on film-substrate system.

In our research, we focus on the extraction of initial shear modulus and only consider neo-Hookean constitutive materials. Here, the materials are assumed incompressible. The aspiration length l is a function of the following parameters:

$$l = f(p, t, R, \mu_f, \mu_s) \quad (1)$$

Appling dimensional analysis and Buckingham Pi theorem [21] on Eq.(1), we get

$$\frac{P}{\mu_s} = \prod \left(\frac{\mu_f}{\mu_s}, \frac{l}{R}, \frac{t}{R} \right) \quad (2)$$

Where \prod is a dimensionless function. The relationship between p/μ_s and l/R can be obtained through the aspiration experiment. Meanwhile, the film thickness and the shear modulus of the substrate are easy to measure. From the data above, the film-substrate modulus ratio μ_f/μ_s can be extracted, and then the film modulus μ_f can be obtained.

2.2 FEM modeling and data collection

We conduct finite element simulations on the pipette aspiration test by using the commercial software ABAQUS 6.14 under the guideline of dimensional analysis above.

Using the Python batch processing in ABAQUS, we set different film thickness t and different film substrate modulus ratio μ_s/μ_f . Zhao et al.[12]analyzed the film-substrate system aspiration test and concluded that when $t/R > 2$, the modulus of film can be measured directly without considering the substrate effect. The calculation deviation will increase when the film becomes thinner. We set t/R from 0.1 to 2, since during the experiment, the pipe radius R can be adjusted to make sure t/R in a reasonable range. Meanwhile, due to the film-substrate materials in vivo have softer substrates and harder films, such as cells and visceral tissues. Therefore, we consider the range of μ_s/μ_f between 1 and 10.

Based previous research [11], we select the data with $t/R \leq 0.3$ for extracting the initial shear modulus, and use interpolation to obtain 100 points on the $p/\mu_s - t/R$ curve uniformly. In this way, the calculated data set is divided into two categories: one is the relationship between p/μ_s and t/R , and the film thickness ratio to the inner diameter of the pipe t/R ; the other one is the film-substrate modulus ratio μ_f/μ_s . These two categories correspond to the independent variable X and the dependent variable y respectively. For each curve, since its t/R data points are fixed, its 100 p/μ_s data points can fully represent the information of the curve. Therefore, the X of each example is a 101-dimensional vector composed of 100 p/μ_s data points and t/R of the example. Thus, this problem has 101-dimensional features. The X of the training set can be written as follows:

$$X = \begin{bmatrix} \left(\frac{p}{\mu_s}\right)_1^{(1)} & \dots & \left(\frac{p}{\mu_s}\right)_{100}^{(1)} & \left(\frac{t}{R}\right)^{(1)} \\ \left(\frac{p}{\mu_s}\right)_2^{(1)} & \dots & \left(\frac{p}{\mu_s}\right)_{100}^{(2)} & \left(\frac{t}{R}\right)^{(2)} \\ \vdots & \ddots & \vdots & \vdots \\ \left(\frac{p}{\mu_s}\right)_m^{(1)} & \dots & \left(\frac{p}{\mu_s}\right)_{100}^{(m)} & \left(\frac{t}{R}\right)^{(m)} \end{bmatrix}, m = 4186 \quad (3)$$

The label (y) of this question is μ_f/μ_s . Machine learning models are established to depict the relationship between the independent variable and the dependent variable . The data set has a total of 4186 training sets and 288 test sets. We make the data points of the test set distribute uniformly for different t/R and μ_f/μ_s to see the parameter effect on errors.

3 Machine learning models

Machine learning methods are widely applied in recent researches, because these methods have the powerful non-linear fitting ability. Here, we focus on two popular kinds of machine learning methods: deep neural network (DNN), and tree-based ensemble learning (XGBoost).

We build our deep neural network model under Pytorch, the popular deep learning framework among researchers. After optimization, our DNN model consists of 5 hidden layers, shown in Figure 2. The activation function of each layer is rectified linear unit (ReLU). To avoid over-fitting, we use dropout method to randomly disconnect the links of some neurons to achieve regularization.

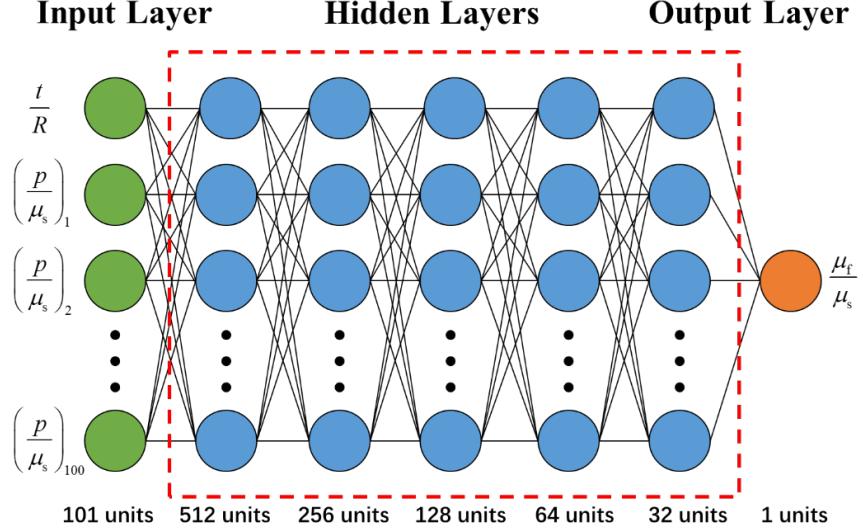


Figure. 2 Schematic of the deep neural network model

The loss function of the neural network model we choose is the mean square error (MSE), which is denoted as

$$MSE = \frac{1}{n} \sum_{i=1}^m (y_i - \hat{y})^2 \quad (4)$$

XGBoost is one of the most popular ensemble learning methods consisted of decision trees. XGBoost algorithm is based on the thought of boosting. In this paper, we use this method to build our predicting model with the training data in Section 2.2. During the modeling, sklearn and xgboost, python packages, are applied.

4 Results

To verify the predictive ability of our explicit model, test set is used to conduct the validation. Here, we define the error as

$$error = \left| \left(\frac{\mu_f}{\mu_s} \right)^{pred} - \left(\frac{\mu_f}{\mu_s} \right)^{act} \right| / \left(\frac{\mu_f}{\mu_s} \right)^{act} \quad (5)$$

Machine learning models	DNN	XGBoost
Training set error	1.26%	0.43%
Test set error	1.83%	1.79%

Table 1: Average percentage errors of training set and test set for the machine learning models (DNN, XGBoost).

The superscript “pred” stands for the predictive value, and the superscript “act” stands for the actual value. The average errors of the two machine learning models are listed in Table 1. We further demonstrate the error frequency distributions of the two models in Figure 3. Both of the models have excellent overall predictive and generalization abilities. In general, the errors of DNN are more concentrated in the range of less than 1%.

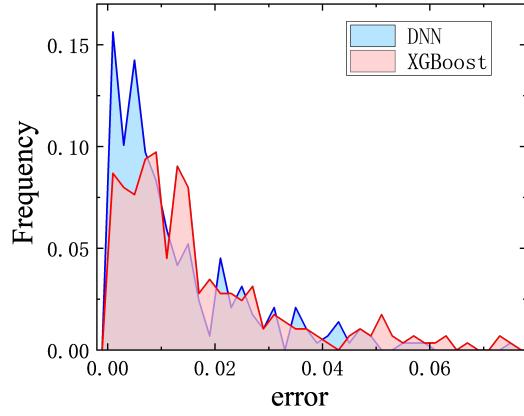
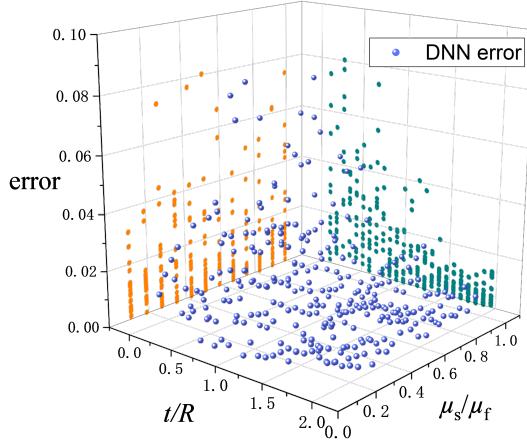
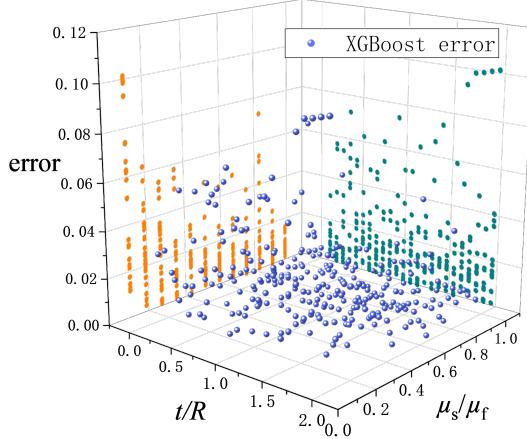


Figure. 3 The error frequency distributions of the test set of DNN and XGBoost models

We further demonstrate the parameter effects on errors, and show the results in Figure 4. The errors of DNN are more sensitive to the parameter t/R than those in XGBoost model. When t/R is small, its errors have a tendency to increase; while the errors of XGBoost are more uniform distributed in the entire range of values.



(a)



(b)

Figure. 4 The error distributions for different parameters. (a) DNN model, (b) XGBoost model.

5 Discussions

Machine learning methods are applied to establish high predictive power models of the pipette aspiration test of film-substrate soft materials. With the results presented in Section 4, we have the following discussions.

First, through dimensional analysis and applying Pi theorem in Section 2.1, we transform Eq (1) into Eq (2). During this process, we successfully reduce the parameters of the problem, which is useful to reduce the computational pressures and the training time. For problems with more complex physical parameters,

dimensional analysis can be a powerful tool to implement feature reduction.

Second, as the average errors of the training set and test set presented in Table 1, all our machine learning models have high predictive and generalization ability for the pipette aspiration test on film-substrate soft materials. As mentioned in section 3.2, smaller t/R leading larger error in the DNN case, while the XGBoost method is less sensitive to this parameter. However, by implement ANOVA (0.1 significant level), we find there is no significant difference in performance between DNN and XGBoost when $t/R < 0.5$. Therefore, the XGBoost and DNN methods are arguably the same suitable for this kind of problem, which collects loading curves and some other parameters and predicts the mechanical properties of soft materials.

Third, we also applied CNN and RNN models, but these models have even worse performance than simple DNN models. The test errors of CNN models are about 10 percent, while RNN models' are around 8 percent. After carefully checking, we decide to use DNN and Xgboost as our models.

Finally, through feature-importance analysis with XGBoost, we have a bit more insight into our problem. The results are shown in Figure 5. We have two main conclusions from the feature-importance analysis. First, feature 100 representing the parameter t/R is one of the most important features, since t/R is the significant variable of the dimensionless function Eq (2). Second, features with a relatively smaller or larger index (0-20, 90-100) occupy over 90% of importance among all features. Except feature 100, other features stand for the start points and endpoints of our loading curves, which indicates we can determine the loading curves just with these points. Therefore, the second conclusion can also prove Eq. (5) we proposed in section 2.3. In our pipette aspiration test, we should focus more on the start points and endpoints of the loading curves. Meanwhile, we can reduce the features of the middle part of the curves to improve training efficiency.

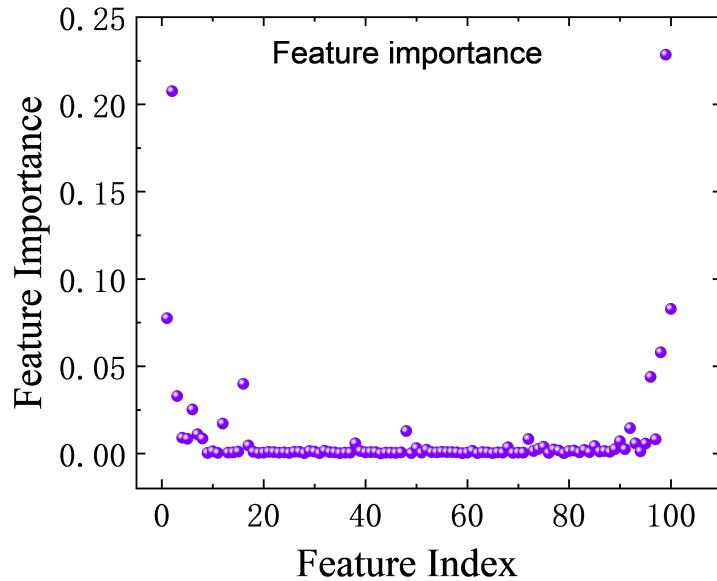


Figure. 5 Feature importance of all features calculated by XGBoost.

6 Conclusions

In this paper, we propose a statistical learning framework for establishing pipette aspiration method for characterizing mechanical properties of film-substrate soft materials. Towards this goal, dimensional analysis, FEM simulations, and machine learning methods are carried out.

Combined with dimensional analysis and large-scale finite element simulation, a larger-scale data set of this problem can be obtained. Machine learning models (deep neural network and XGBoost) are implemented to build high predictive models. With ANOVA, we further verify that DNN and XGBoost methods have the similar predictive power for this problem. Feature importance analysis is also applied to determine the more important features.

7 Source code

<https://github.com/SaintAugustus/Comp-576>

References

- [1] Hochmuth RM. Micropipette aspiration of living cells. *Journal of Biomechanics*. 2000;33:15-22.
- [2] Aoki T, Ohashi T, Matsumoto T, Sato M. The pipette aspiration applied to the local stiffness measurement of soft tissues. *Annals of Biomedical Engineering*. 1997;25:581-7.
- [3] Boudou T, Ohayon J, Arntz Y, Finet G, Picart C, Tracqui P. An extended modeling of the micropipette aspiration experiment for the characterization of the Young's modulus and Poisson's ratio of adherent thin biological samples: Numerical and experimental studies. *Journal of Biomechanics*. 2006;39:1677-85.
- [4] Alexopoulos LG, Haider MA, Vail TP, Guilak F. Alterations in the mechanical properties of the human chondrocyte pericellular matrix with osteoarthritis. *Journal of Biomechanical Engineering-Transactions of the Asme*. 2003;125:323-33.
- [5] Kauer M, Vuskovic V, Dual J, Szekely G, Bajka M. Inverse finite element characterization of soft tissues. *Medical Image Analysis*. 2002;6:275-87.
- [6] Wilder EA, Guo S, Lin-Gibson S, Fasolka MJ, Stafford CM. Measuring the modulus of soft polymer networks via a buckling-based metrology. *Macromolecules*. 2006;39:4138-43.
- [7] Huang J, Juszkiewicz M, de Jeu WH, Cerda E, Emrick T, Menon N, et al. Capillary wrinkling of floating thin polymer films. *Science*. 2007;317:650-3.
- [8] Huang J, Juszkiewicz M, de Jeu WH, Cerda E, Emrick T, Menon N, et al. Capillary wrinkling of floating thin polymer films. *Science*. 2007;317:650-3.
- [9] Cao YP, Zheng XP, Li B, Feng XQ. Determination of the elastic modulus of micro- and nanowires/tubes using a buckling-based metrology. *Scripta Materialia*. 2009;61:1044-7.
- [10] Stafford CM, Harrison C, Beers KL, Karim A, Amis EJ, Vanlandingham MR, et al. A buckling-based metrology for measuring the elastic moduli of polymeric thin films. *Nature Materials*. 2004;3:545-50.

- [11] Zhang MG, Cao YP, Li GY, Feng XQ. Pipette aspiration of hyperelastic compliant materials: Theoretical analysis, simulations and experiments. *Journal of the Mechanics and Physics of Solids*. 2014;68:179-96.
- [12] Zhao RG, Sider KL, Simmons CA. Measurement of layer-specific mechanical properties in multilayered biomaterials by micropipette aspiration. *Acta Biomaterialia*. 2011;7:1220-7.
- [13] Pellicer-Valero OJ, Ruperez MJ, Martinez-Sanchis S, Martin-Guerrero JD. Real-time biomechanical modeling of the liver using Machine Learning models trained on Finite Element Method simulations. *Expert Systems with Applications*. 2020;143:113-83.
- [14] Alekya B, Rao S, Pandya HJ. Engineering approaches for characterizing soft tissue mechanical properties: A review. *Clinical Biomechanics*. 2019;69:127-40.
- [15] Li X, Liu ZL, Cui SQ, Luo CC, Li CF, Zhuang Z. Predicting the effective mechanical property of heterogeneous materials by image based modeling and deep learning. *Computer Methods in Applied Mechanics and Engineering*. 2019;347:735-53
- [16] Ford E, Maneparambil K, Rajan S, Neithalath N. Machine learning-based accelerated property prediction of two-phase materials using microstructural descriptors and finite element analysis. *Computational Materials Science*. 2021;191:110-7.
- [17] Liang L, Liu ML, Sun W. A deep learning approach to estimate chemically-treated collagenous tissue nonlinear anisotropic stress-strain responses from microscopy images. *Acta Biomaterialia*. 2017;63:227-35.
- [18] Yu Z, Ye S, Sun YL, Zhao H, Feng XQ. Deep learning method for predicting the mechanical properties of aluminum alloys with small data sets. *Materials Today Communications*. 2021;28:102-570.
- [19] Ye S, Li B, Li QY, Zhao HP, Feng XQ. Deep neural network method for predicting the mechanical properties of composites. *Applied Physics Letters*. 2019;115:161-16.
- [20] Chen TQ, Guestrin C, Assoc Comp M. XGBoost: A Scalable Tree Boosting System. 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD). San Francisco, CA2016. p. 785-94.
- [21] Barenblatt GI, Barenblatt GI, Isaakovich BG. Scaling, self-similarity, and intermediate asymptotics: dimensional analysis and intermediate asymptotics: Cambridge University Press; 1996.