

Learning the Implicit Strain Reconstruction in Ultrasound Elastography Using Privileged Information

Zhifan Gao^a, Sitong Wu^{b,c}, Zhi Liu^d, Jianwen Luo^d, Heye Zhang^{e,*},
Mingming Gong^f, Shuo Li^{a,g,h,*}

^aDepartment of Medical Biophysics, Western University, London N6A 3K7, Canada

^bShenzhen Institutes of Advanced Technology, Chinese Academy of Sciences,
Shenzhen 518055, China

^cShenzhen College of Advanced Technology, University of Chinese Academy of Sciences,
Shenzhen 518055, China

^dDepartment of Biomedical Engineering, School of Medicine, Tsinghua University,
Beijing 100084, China

^eSchool of Biomedical Engineering, Sun Yat-Sen University, Shenzhen 510006, China

^fSchool of Mathematics and Statistics, The University of Melbourne,
Parkville, VIC 3010, Australia

^gDepartment of Medical Imaging, Western University, London N6A 3K7, Canada

^hDigital Imaging Group of London, London N6A 3K7, Canada

Abstract

Quasi-static ultrasound elastography is an importance imaging technology to assess the conditions of various diseases through reconstructing the tissue strain from radio frequency data. State-of-the-art strain reconstruction techniques suffer from the inexperienced user unfriendliness, high model bias, and low effectiveness-to-efficiency ratio. The three challenges result from the explicitness characteristic (i.e. explicit formulation of the reconstruction model) in these techniques. For these challenges, we are the first to develop an *implicit* strain reconstruction framework by a deep neural network architecture. However, the classic neural network methods are unsuitable to the strain reconstruction task because they are difficult to impose any direct influence on the intermediate state of the learning process. This may lead the map learned by the neural network to be biased with the desired map. In order to correct the intermediate state of the learning process, our frame-

*Corresponding author: Heye Zhang (zhangheye@mail.sysu.edu.cn) and Shuo Li (slishuo@gmail.com)

work proposes the learning-using-privileged-information (LUPI) paradigm with causality in the network. It provides the causal privileged information besides the training examples to help the network learning, while makes these privileged information unavailable at the test stage. This improvement can narrow the search region of the map learned by the network, and thus prompts the network to evolve towards the actual ultrasound elastography process. Moreover, in order to ensure the causality in LUPI, our framework proposes a physically-based data generation strategy to produce the triplets of privileged information, training examples and labels. This data generation process can approximately describes the actual ultrasound elastography process by the numerical simulation based on the tissue biomechanics and ultrasound physics. It thus can build the causal relationship between the privileged information and training examples/labels. It can also address the medical data insufficiency problem. The performance of our framework has been validated on 100 simulation data, 42 phantom data and 4 real clinical data by comparing with the ground truth performed by an ultrasound simulation system and four state-of-the-art methods. The experimental results show that our framework is well agreed (average bias is 0.065 for strain reconstruction) with the ground truth, as well as superior to these state-of-the-art methods. These results can demonstrate the effectiveness of our framework in the strain reconstruction.

Keywords: Quasi-static ultrasound elastography, implicit model, deep neural network, learning using privileged information, strain reconstruction

1. Introduction

Quasi-static ultrasound elastography is a crucial ultrasound imaging technology to indicate the conditions of diseases (e.g. inflammation, hypertrophy, edema, and fibrosis) by assessing the tissue stiffness (Zaleska-Dorobisz et al., 2014; Hoerig et al., 2017; Shiina et al., 2015). This imaging technology aims to reconstruct the tissue strain distribution within the suspected lesion area according to the radio frequency (RF) data before and after compressing the target tissue by means of an ultrasound transducer (Ophir et al., 1991). It thus has the ability to evaluate the tissue elasticity (indicated by the reconstructed strain) that is difficultly acquired in the commonly-used B-mode ultrasound imaging (Barr et al., 2015; Ferraioli et al., 2015; Cosgrove et al., 2017; Barr et al., 2017; Cosgrove et al., 2013). Its ability of tissue elasticity

13 evaluation brings in the significant improvement of clinical diagnosis accu-
14 racy for many diseases (e.g. 15.5% accuracy improvement in the diagnosis
15 of malignant lymph node (Zhang et al., 2009), because the elasticity is a
16 significant indicator of organic diseases. European Federation of Societies
17 for Ultrasound in Medicine and Biology (EFSUMB) have shown that the
18 quasi-static ultrasound elastography is able to reach about 90% sensitivity
19 and specificity in disease diagnosis of many organs, such as breast, thyroid
20 and prostate (Bamber et al., 2013).

21 However, the state-of-the-art strain reconstruction techniques in quasi-
22 static ultrasound elastography suffer from three major challenges (see Figure
23 1). All these challenges result from their all use of explicit models to solve
24 the strain reconstruction problem, i.e. the intrinsic solution formulation for
25 strain reconstruction is explicit.

26 The first challenge is the weak model adaptation, i.e. a single model is
27 difficult to adapt all RF data. It originates from the explicit characteristic
28 of the strain reconstruction models in all existing literature (Rivaz et al.,
29 2009, 2011; Brusseau et al., 2008; Kuzmin et al., 2015; Yuan et al., 2011;
30 Pellot-Barakat et al., 2004). The explicitness leads that all models can ob-
31 tain the strain reconstruction solution (i.e. strain field) only from some RF
32 data, depending on their own characteristics (e.g. field smoothness (Hashemi
33 and Rivaz, 2017) and optical flow consistency (Zakaria et al., 2010)). This
34 implies that each explicit model occupies a separate bounded region in the
35 solution space for the strain reconstruction problem. i.e. the solution set
36 of all explicit models is composed of many separate regions. Hence, all ex-
37 plicit models may poorly perform on the RF data whose corresponding strain
38 reconstruction solutions fall outside their solution set. This indicates the ex-
39 plicit models weakly adapts the various RF data, and thus may provide the
40 strain reconstruction solution with high model bias.

41 The second challenge is the contradiction between the effectiveness and ef-
42 ficiency in the explicit methods. The effectiveness denotes the reconstruction
43 performance of the ultrasound elastography method, and the efficiency de-
44 notes its computational cost. Although having high computational efficiency,
45 the methods with simple formulated model are difficult to accurately recon-
46 struct the tissue strain in noise-corrupted RF data. This is because these
47 methods may be apt to produce the noise-induced error (e.g. peak error and
48 jitter error) (Walker and Trahey, 1995). To reduce the noise disturbance,
49 varieties of complex reconstruction model (e.g. biomechanical model) are
50 developed at the cost of the computational efficiency due to the extensive

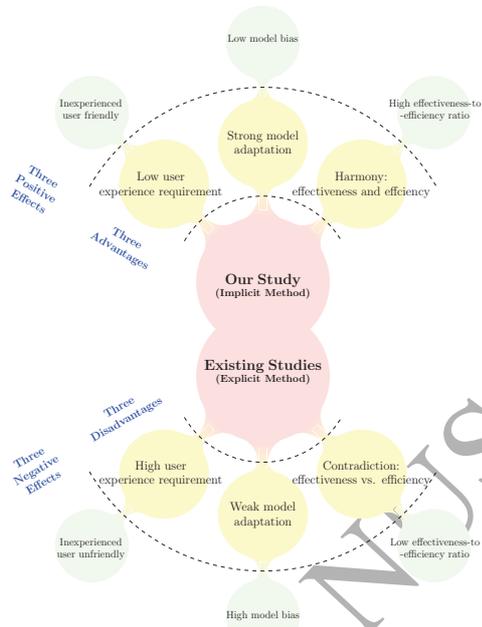


Figure 1: Superiority of our study to the previous quasi-static ultrasound elastography technologies. The previous studies aim at the explicit strain reconstruction method, and thus bring in three disadvantages (dependency on user’s experience, weak model adaptation, and effectiveness-efficiency contradiction). Our study is the first to focus on the implicit method. It can overcome the above disadvantages and bring the three benefits (inexperienced use friendly, low model bias and high effectiveness-to-efficiency ratio) to the quasi-static ultrasound elastography.

51 iterative calculations (Lu et al., 2014). Thus, the existing explicit methods
 52 have low effectiveness-to-efficiency ratio, i.e. the ratio of the reconstruction
 53 performance to the computational cost of the ultrasound elastography
 54 method.

55 The final challenge is the requirement of the manual model selection when
 56 using the explicit methods. The manual model selection comes from the dif-
 57 ferent advantages and disadvantages of various explicit methods. To satisfy
 58 the requirement of the specific clinical scenario, the user should select the
 59 appropriate one in various explicit methods for strain reconstruction accord-
 60 ing to data characteristics and performance requirement. Thus, the manual
 61 model selection highly relies on the user’s professional experience and is un-
 62 friendly to the inexperienced user.

63 Classic implicit methods (i.e. the general model without explicitly mod-
64 eling the process within the input-output map) are unsuitable for strain re-
65 construction owing to its time-consuming characteristic, although they can
66 be naturally considered as a potential alternative to tackle the challenges in
67 the explicit method. All these methods (e.g. variational inference (Wain-
68 wright and Jordan, 2008) and particle filtering (Fearnhead, 2002)) require
69 numerous computations in sampling from the approximate data density dur-
70 ing the optimization (e.g. Monte Carlo methods (Gilks et al., 1995)). The
71 high computational requirement limits these implicit methods for strain re-
72 construction in clinical practice. This is the potential reason that no implicit
73 method for strain reconstruction has appeared in the literature to our best
74 knowledge.

75 Recent implicit methods based on the deep neural network have the po-
76 tential to address the challenges of the explicit methods and the classic im-
77 plicit methods for strain reconstruction. First, these methods can largely
78 reduce the computation with respect to the classic methods due to no itera-
79 tion calculation. Then, they are able to cope with the weak model adaptation
80 challenge and lead to the low model bias by learning the reconstruction pro-
81 cess from a large set of solution space. Accordingly, they can naturally ad-
82 dress the effectiveness-efficiency contradiction challenge because of their low
83 model bias and high computational efficiency. Finally, they are capable to
84 tackle the manual model selection challenge as its ability to well approximate
85 various models from its forward multi-layer inference structure (Schmidhu-
86 ber, 2015; Guo et al., 2016).

87 Nevertheless, the traditional deep neural network is unsuitable for the
88 direct use in the strain reconstruction task. This is because it is incapable to
89 impose any direct influence on the internal feature representation of the input
90 that the network should focus on during the learning process (Shi and Kim,
91 2017; Bisla and Choromanska, 2018). This drawback is particularly serious
92 in the medical fields (such as ultrasound elastography) which require the
93 complex learning models but have limited training data. Thus, the network
94 has the potential to learn the inference model mismatching the actual map
95 between training data and labels, because its internal feature representation
96 may not be related to the target task (Bisla and Choromanska, 2018).

97 In this paper, we propose a brand-new quasi-static ultrasound elastog-
98 raphy framework based on the deep neural network for implicit strain re-
99 construction, i.e. inference the map from the RF data to issue strain field
100 (RF-to-strain map). For enabling the ability of the network to directly influ-

101 ence the internal feature representation, our framework applies an advanced
 102 learning paradigm named learning using privileged information (LUPI) (Vap-
 103 nik and Izmailov, 2015) to the network, and is thus called UE-LUPI. The
 104 UE-LUPI provides some additional information called privileged information
 105 besides the training example to help the network learning, while makes these
 106 privileged information unavailable at the test stage (Vapnik and Vashist,
 107 2009). It is able to narrow the search region of the map learned by the net-
 108 work, and thus prompts the network to evolve towards the actual ultrasound
 109 elastography process. Specifically, we consider the tissue displacement as the
 110 appropriate privileged information. According to the quasi-static ultrasound
 111 elastography (Ophir et al., 1991; Treece et al., 2011), the RF data is used
 112 to compute the tissue displacement, and the tissue displacement is used to
 113 compute the tissue strain. It brings in the causal relationship: RF data \rightarrow
 114 tissue displacement \rightarrow tissue strain. This causal relationship can reduce the
 115 model complexity, as well as satisfy the no-harm guarantee (i.e. not harm the
 116 learning process) in the LUPI paradigm (Lambert et al., 2018). Moreover,
 117 this causality endows the network with the ability to influence the interme-
 118 diate state (i.e. displacement) in the learning process, and thus induce the
 119 network to gradually learn the actual ultrasound elastography process. How-
 120 ever, the actual tissue displacement and strain tags can not obtained from RF
 121 data by any direct measurement technique in clinical practice. For tackling
 122 this problem, UE-LUPI proposes a strategy to generate the triplet of train-
 123 ing data, privileged information and training labels based on the numerical
 124 simulation according to tissue biomechanics and ultrasound physics. This
 125 strategy is capable to provide reliable privileged information by building the
 126 causal relationship among training data, privileged information and training
 127 labels to approximately describe the actual ultrasound elastography process.
 128 It is motivated by the extensive use of numerical simulation in the literature
 129 of quasi-static ultrasound elastography (e.g. (Rivaz et al., 2011; Hashemi
 130 and Rivaz, 2017)). In addition, this strategy is able to address the clinical
 131 data insufficiency issue through generating various training data and labels.
 132 It can thus overcome the model overfitting, as well as preserve all rare and
 133 precious clinical data for the model validation.

134 Our contributions can be summarized as follows:

- 135 1. We are the first to propose a framework (UE-LUPI) for the implicit
 136 strain reconstruction in order to improve the user friendliness, model adap-
 137 tation, and effectiveness-to-efficiency ratio in the current clinical application.
- 138 2. We use the causal LUPI learning paradigm to help the network in-

139 ference by correcting the intermediate state in the learning process. This
140 strategy can prompt the RF-to-strain map learned by the network to evolve
141 towards the actual ultrasound elastography process. Particularly, the tissue
142 displacement is used as the causal privileged information for improving the
143 performance of the RF-to-strain inference.

144 we consider the tissue displacement as the appropriate privileged infor-
145 mation. The causal relationship of the tissue displacement with the RF data
146 and tissue strain can reduce the model complexity, as well as satisfy the
147 no-harm guarantee

148 3. We propose a strategy based on tissue biomechanics and ultrasound
149 physics to obtain the privileged information that is unavailable in clinical
150 practice. This strategy can lead that the data generation process to ap-
151 propriately approximate the actual ultrasound elastography process. It can
152 ensure the causality of the privileged information (tissue displacement) with
153 the training examples (RF data) and labels (tissue strain), and cause the net-
154 work learning with the privileged information to be suitable for real scenarios.
155 It is also capable to address the model overfitting due to data insufficiency
156 and keep the rare clinical data for model validation.

157 This study has advantages over our preliminary study in MICCAI con-
158 ference (Wu et al., 2018): (1) applying the LUPI learning paradigm to the
159 neural network for correcting the intermediate state in the network learning.
160 The LUPI requires the additional information (called privileged information)
161 besides the training examples and labels in the training stage. Our study has
162 a LUPI learning paradigm by bringing in the displacement field as the priv-
163 ileged information to supervise the intermediate layer in the training stage,
164 while the preliminary study applies two cascade networks, both of which
165 have no privileged information. (2) adding the causality among the RF data,
166 tissue displacement and tissue strain to learn an end-to-end RF-strain map
167 which do not require the tissue displacement in the test stage. This causal-
168 ity and the conditional independence between the input and output of the
169 reconstruction model on the privileged information results in the reduction
170 of the upper bound of the generalization error. (3) enlarging the experimen-
171 tal datasets and comparing with more state-of-the-art methods for better
172 validating the effectiveness and superiority of our framework.

173 **2. Related Work**

174 *2.1. LUPI for Influencing the Internal Feature Representation*

175 LUPI is an advanced paradigm inspired by human experience of learn-
 176 ing with teacher (Vapnik and Izmailov, 2015; Sharmanska et al., 2014): it is
 177 better that the teacher provide intuitive comments, comparisons, and expla-
 178 nations to students than simply tell them the answer (Lambert et al., 2018).
 179 Accordingly, the teacher in LUPI supplies the student with additional in-
 180 formation (called privileged information) besides the training examples and
 181 labels only available in the training stage (Pechyony and Vapnik, 2010). The
 182 LUPI paradigm can be simply described as (Pechyony and Vapnik, 2010;
 183 Vapnik and Izmailov, 2015):

$$\begin{aligned} &\text{given } \{x_i, z_i, y_i\}, \quad i = 1, \dots, N, x_i \in X, z_i \in Z, y_i \in Y \\ &\text{purpose } \text{learn } X \rightarrow Y \text{ with the small generalization error} \end{aligned} \quad (1)$$

184 where N is the number of training samples, x_i , z_i and y_i are training ex-
 185 amples, privileged information, training labels, respectively. As a general
 186 framework, LUPI has shown its success in varieties of learning tasks (Fey-
 187 ereisl et al., 2014; Lapin et al., 2014; Lambert et al., 2018; Shi and Kim, 2017;
 188 Duan et al., 2014; Zheng et al., 2016). In the neural network, the privileged
 189 information in LUPI can be turned into various forms for specific tasks, for
 190 example, the statistical constraints in the autoencoders (Hjelm et al., 2019),
 191 the task feedback in the multi-task learning (Thung and Wee, 2018; Zhang
 192 and Yang, 2018), the label transformation in the deep supervision (Lee et al.,
 193 2015; Dong et al., 2018), the bounding annotation in the image classification
 194 (Sharmanska et al., 2013), the filter-based statistical structure in the tex-
 195 ture synthesis (Gatys et al., 2015), and the multiple inputs in the multi-view
 196 learning (Shi et al., 2018; Meng et al., 2018).

197 The main superiority of our LUPI paradigm is that the proposed priv-
 198 ileged information has *causality* with the training examples and labels, in
 199 contrast to the *dependency* in the existing studies. The proposed *causal*
 200 *LUPI* is defined as the LUPI paradigm whose training examples, privileged
 201 information, training labels have the causal relationship, i.e. X causes Z ,
 202 and Z causes Y . In our framework, the influence of the internal feature
 203 representation by privileged information can be viewed as the correlation
 204 of the intermediate state in the learning process, which is distinct from the
 205 commonly-used latent space constraint schemes (Chang et al., 2018; Engel
 206 et al., 2018). This is because the constraints aims to ameliorate the certain

207 characteristics of the latent space, rather than provide the explicit knowledge
 208 about the intermediate state of the learned map $X \rightarrow Y$. This difference
 209 brings in the superiority of the proposed UE-LUPI that the learned map can
 210 evolve towards the actual ultrasound elastography process, because the inter-
 211 mediate state of model inference can be gradually corrected by the privileged
 212 information during the network optimization.

213 *2.2. Quasi-Static Ultrasound Elastography*

214 Most quasi-static ultrasound elastography methods fall into two main
 215 categories: gradient-based methods and direct methods. The gradient-based
 216 methods usually firstly compute the time delay or phase shift between the
 217 pre- & post-compression RF data to estimate the tissue displacement. Then
 218 they use the spatial derivation to reconstruct the strain field (Zahiri-Azar
 219 and Salcudean, 2006; Ophir et al., 1991; Omidyeganeh et al., 2017; Yuan and
 220 Pedersen, 2015; Zhu and Hall, 2002). The direct methods reconstruct the
 221 strain directly from the RF data (Hasan et al., 2012). Varieties of methods
 222 aim to reduce the error from the noise disturbance and signal decorrelation,
 223 such as dynamic programming (Rivaz et al., 2009, 2014), least square method
 224 (Jiang and Hall, 2015; Shaswary et al., 2016), Kalman filter (Rivaz et al.,
 225 2011), and non-temporal analysis (DiBattista and Noble, 2014; Techavipoo
 226 and Varghese, 2004).

227 Few previous studies used the neural network or other machine learning
 228 methods in ultrasound elastography. Amidabadi et al. (2018) used support
 229 vector machine to classify peak hopping and jitter error for improving the
 230 performance of time delay estimation. Hoerig et al. (2017) modeled the
 231 force-displacement relationship in ultrasound elastography by the finite el-
 232 ement analysis whose computational complexity and model complexity are
 233 reduced by the neural network. Kibria and Rivaz (2018) applied the neural
 234 network to compute the optical flow between RF scans for tissue displace-
 235 ment estimation.

236 The main difference of our UE-LUPI with the existing ultrasound elastog-
 237 raphy studies (including the neural-network-based methods) is that we are
 238 the first to learn the RF-to-strain map by the neural network for implicitly
 239 reconstructing the tissue strain. This leads to the superior characteristics of
 240 UE-LUPI including the improvement of user friendliness, the model adapta-
 241 tion, and the effectiveness-to-efficiency ratio.

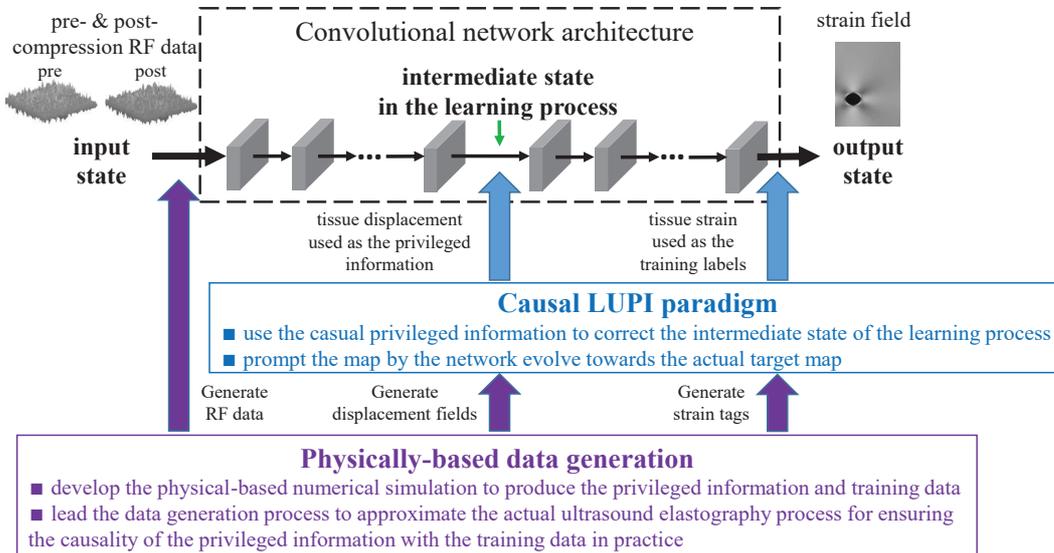


Figure 2: Overview of UE-LUPI. UE-LUPI proposes an implicit framework based on the deep neural network to reconstruct the strain field from pre- & post-compression 2D RF data. In UE-LUPI, a causal learning-using-privileged-information (LUPI) learning paradigm is proposed to correct the intermediate state in the network learning process. Then, a physically-based data generation is proposed to produce the privileged information (tissue displacement) and training data (RF data and tissue strain) for ensuring the causality of the privileged information.

242 3. Methodology

243 Our newly proposed UE-LUPI provides an implicit framework by the deep
 244 neural network to construct the tissue strain from RF data, shown in Figure
 245 2. UE-LUPI enables the LUPI learning paradigm in strain reconstruction by
 246 using the tissue displacement as the privileged information. This selection
 247 brings in the causal relationship of the privileged information with the train-
 248 ing examples (i.e. the RF data) and labels (i.e. the tissue strain). It thus
 249 causes the training examples and labels have the conditional independency
 250 according to the ultrasound elastography principle. As regards this causality,
 251 the privileged information is incorporated into a specially-designed loss term
 252 of the neural network for correcting the intermediate state of the learning
 253 process (detailed in Section 3.1 and Figure 4) This strategy enables the di-
 254 rect guidance of the inner feature representation of the network for correcting
 255 the intermediate state in the learning process, which is superior to the clas-

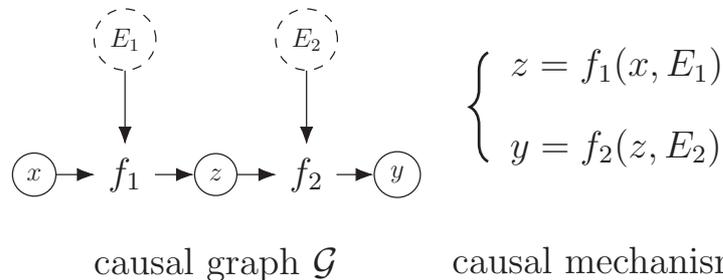


Figure 3: The functional causal model to show the direct causal relationship in the proposed causal LUPI learning paradigm.

256 sic neural network methods (i.e. without the LUPI paradigm). Thus, it can
 257 narrow the search region of the map learned by the network and enriched the
 258 feedback information. It further leads that the map learned by UE-LUPI is
 259 able to evolve towards the actual ultrasound physical process. Furthermore,
 260 for ensuring the causality of the privileged information, UE-LUPI applies the
 261 numerical simulation to produce the triplets of RF data, tissue displacement
 262 and tissue strain according to tissue biomechanics and ultrasound physics
 263 (detailed in Section 3.2 and Figure 5). In addition, this strategy enables
 264 the training process of the neural network in the absence of clinical tissue
 265 displacement and strain tags. The above behaviors lead the effectiveness of
 266 the UE-LUPI in the strain reconstruction task.

267 3.1. Causal privileged information for correcting intermediate state in the 268 network learning process

269 **Formulation of the causal LUPI learning paradigm.** Our UE-LUPI
 270 has the causal LUPI learning paradigm by considering tissue displacement
 271 as the privileged information, i.e. add a causal constraint in Equation (1).
 272 It can be formulated as

$$\begin{array}{ll}
 \text{given} & \{x_i, z_i, y_i\}, \quad i = 1, \dots, N, x_i \in X, z_i \in Z, y_i \in Y \\
 \text{subject to} & X \rightarrow Z \rightarrow Y \text{ has the causal relationship} \\
 \text{purpose} & \text{learn } X \rightarrow Y \text{ with the small generalization error}
 \end{array} \quad (2)$$

273 where N is the number of training samples, x_i , z_i and y_i are the training
 274 examples, privileged information, training labels, respectively.

275 The causal relationship in Equation (2) can be formulated by the func-
 276 tional causal model (Goudet et al., 2018; Zhang et al., 2016):

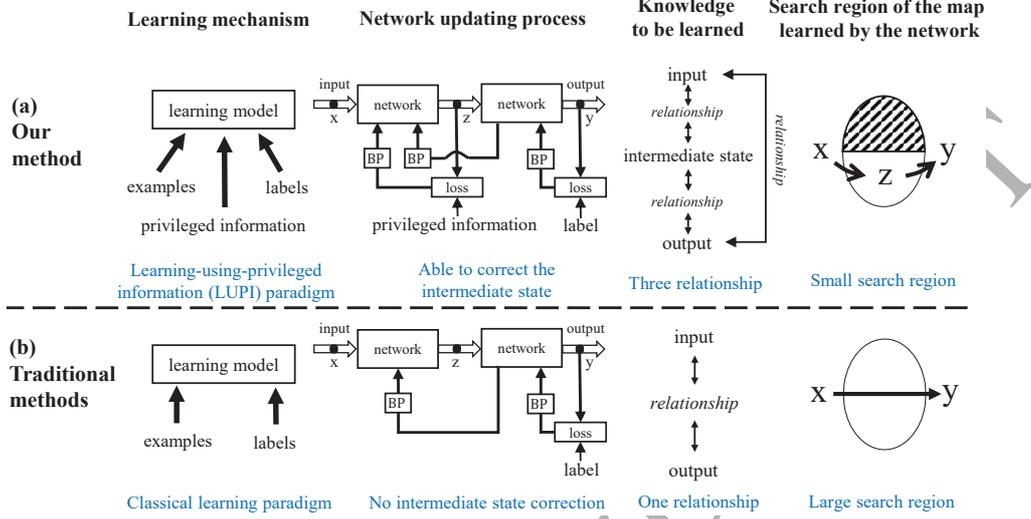


Figure 4: Superiority of UE-LUPI to the traditional deep neural network methods. (a) Our UE-LUPI applies the learning-using-privileged-information learning paradigm. It can correct the intermediate state in the network learning process, and leads the network to learn from more knowledge. Thus, it can narrow the search region of the map learned by the network, and prompt this map to evolve towards the actual ultrasound elastography process. (b) Traditional neural network methods cannot influence the intermediate state in the learning process, and thus may lead the learned map to be biased with the actual target map.

$$S_j \leftarrow f_j(S_{Pa(k;G)}, E_k) \quad (3)$$

277 where $\mathbf{S} = \{S_1, S_2, \dots\}$ is the set of observation variables in the causal graph
 278 \mathcal{G} . $S_{Pa(k;G)}$ is set of the direct causal variable S_j based on the causal mech-
 279 nism f_j . E_1 and E_2 are the noise. In the proposed LUPI learning paradigm,
 280 the casual graph \mathcal{G} is specified as the Equation (4) and shown in Figure 3.

$$Z = f_1(X, E_1), \quad Y = f_2(Z, E_2) \quad (4)$$

281 Thus, the purpose to learn the $X \rightarrow Y$ in Equation (2) can be simplified to
 282 learn the two maps $f_1 : X \rightarrow Z$ and $f_2 : Z \rightarrow Y$.

283 **Supervision-enriched neural network to solve the causal LUPI**
 284 **problem.** In order to solve the causal LUPI problem, UE-LUPI uses the
 285 neural network to learn the functions f_1 and f_2 . In particular, UE-LUPI
 286 constructs a loss term to provide the feedback information from the privi-
 287 leged information to the intermediate layer of the network (see Figure 4(a)).

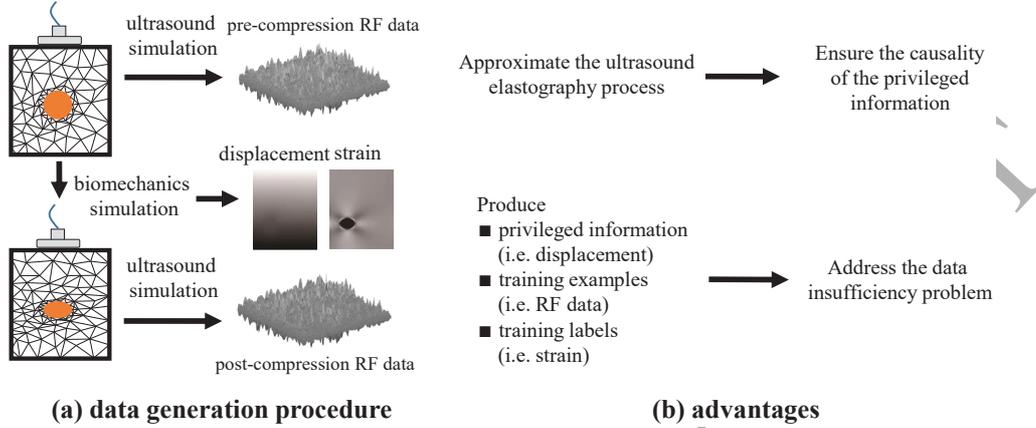


Figure 5: Physically-based data generation in UE-LUPI. (a) The tissue displacement and strain are computed by the biomechanics simulation of the tissue deformation before and after the compression. The tissue deformation are then used as the digital phantoms to compute the pre- and post-compression RF data by the ultrasound simulation. (b) Two advantages of this data generation strategy: 1) ensure the causality of the privileged information with the training examples and labels in practice by approximating the ultrasound elastography process; 2) address the medical data insufficiency problem, and keep the rare clinical data for model validation rather than consuming them in the model training.

288 With this loss term, the traditional loss formulation of the end-to-end neural
 289 network can be transformed into

$$loss = L(\hat{y}, y) + L(\hat{z}, z) \quad (5)$$

290 where x is the training examples used as the network input. $\hat{y} = f_2(f_1(x))$ is
 291 the predict value of the training labels y in the neural network. $\hat{z} = f_1(x)$ is
 292 the predict value of the privileged information z . L is the loss function. f_1 and
 293 f_2 denote the function map of $X \rightarrow Z$ and $Z \rightarrow Y$, which corresponds to the
 294 front part and back part of the network, respectively. Accordingly, the neural
 295 network in our UE-LUPI aims to learn the map $X \rightarrow Z \rightarrow Y$, while the
 296 classic neural network focuses on the map $X \rightarrow Y$. Their difference is shown
 297 in Figure 4. Thus, the causal privileged information z is able to facilitate
 298 the neural network learning because the extra supervision information are
 299 provided.

300 **Enable to inference the intermediate state of the target map.** Be-
 301 sides using the extra supervision information, UE-LUPI is capable to learn
 302 the intermediate state of the actual target map. According to Equation (2),
 303 the causal privileged information Z contains the knowledge of the interme-

304 diate state of the target map $X \rightarrow Y$, but just having this causal condition
 305 cannot assure that the network learns the actual target map. This is be-
 306 cause there may exist other unknown privileged information which also have
 307 the causality with the training examples and labels, i.e. $\exists X \rightarrow Z^k \rightarrow Y$
 308 ($k = 1, 2, \dots$). This may lead the network evolves towards a biased learn-
 309 ing direction because the single kind of privileged information contains the
 310 partial knowledge of the intermediate state of the target map.

311 Selecting the tissue displacement as the privileged information can ad-
 312 dress the problem of the incomplete intermediate state knowledge. Accord-
 313 ing to the imaging principle of quasi-static ultrasound elastography (Ophir
 314 et al., 1991; Treece et al., 2011), the RF data (i.e. X) is used to compute the
 315 tissue displacement (i.e. Z), and the tissue displacement is used to compute
 316 the tissue strain (i.e. Y). It brings in the causal relationship $X \rightarrow Z \rightarrow Y$.
 317 Then, the tissue displacement is the necessary intermediate observation in
 318 the strain imaging process. Moreover, the tissue displacement can determine
 319 the tissue strain without the help of RF data, which has been validated by
 320 the gradient-based quasi-static methods (detailed in Section 2.2). Thus, the
 321 tissue displacement contains the sufficient information to inference the tissue
 322 strain. This characteristic leads the following proposition:

323 **Proposition 1.** *Denote the random variable of the input, privileged in-*
 324 *formation and output in the learning model with LUPI paradigm as \mathcal{X} , \mathcal{Z} and*
 325 *\mathcal{Y} , respectively. For the UE-LUPI, \mathcal{X} and \mathcal{Y} are conditionally independent*
 326 *on \mathcal{Z} , i.e. $\mathcal{X} \perp \mathcal{Y} | \mathcal{Z}$*

327 The conditional independence in Proposition 1 indicates that we can ob-
 328 serve all information in Y from Z (Koller and Friedman, 2009), i.e. Z con-
 329 tains all information of the relationship between X and Y , and the remaining
 330 information in X not included in Z is not related to Y . Specific in the strain
 331 reconstruction task, the the information in the RF data (X) related to the
 332 tissue strain (Y) are all included in tissue displacement (Z), and the tissue
 333 strain can be computed based on the displacement information without the
 334 need of the RF data. This property can satisfy the no-harm guarantee in the
 335 LUPI paradigm (Lambert et al., 2018), because this privileged information
 336 do not provide the redundant or incomplete knowledge to mislead the learn-
 337 ing model. Accordingly, the privileged information can be constructed into
 338 a loss term to the intermediate layer of the neural network in our UE-LUPI.
 339 The form of the loss term can guide the network training by correcting the

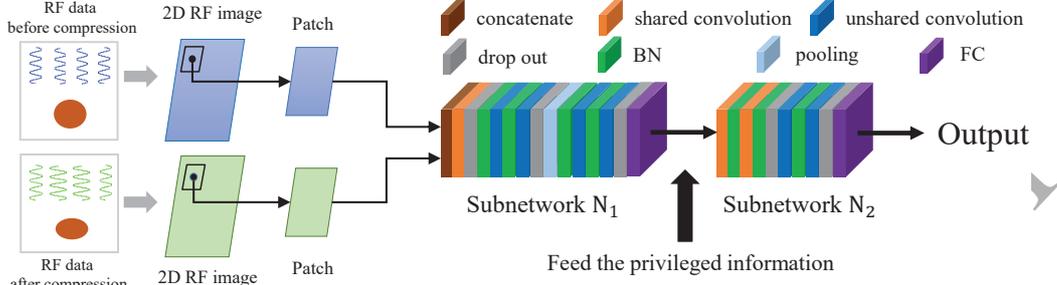


Figure 6: The network details in UE-LUPI.

intermediate state of the learning process, and thus prompt the map learned by the network to evolve towards the actual ultrasound elastography.

Improve the strain reconstruction performance. By the conditional independence in Proposition 1, UE-LUPI can be considered to learn a map by two simple models (i.e. N_1 and N_2) instead of a complex model (i.e. the entire network). This characteristic leads the following proposition:

Proposition 2. *UE-LUPI can reduce the upper bound of the generalization error in the neural network.*

The generalization error of the learning model is the difference between the true error R and training error R_m , where the true error has the following Rademacher generalization bound with probability at least $1 - \delta$ (Bartlett and Mendelson, 2002):

$$R(h) \leq R_m(h) + 4M\mathfrak{R}_m(\mathcal{H}) + M^2 \sqrt{\frac{\log \frac{1}{\delta}}{2m}} \quad (6)$$

where h is the map or function possibly learned by the network in the hypothesis space \mathcal{H} , m is the number of samples, \mathfrak{R}_m is the Rademacher complexity that represents the model complexity. M is the upper bound of infinite norm between h and the true model. δ is a positive marginal constant.

Denote the Rademacher complexity of the traditional neural network as $\mathfrak{R}_m(\mathcal{H})$. In our UE-LUPI, the learning process of the neural network can be considered as consisting of two cascade processes (i.e. $X \rightarrow Z$ and $Z \rightarrow Y$). Thus, the Rademacher complexity of the entire network can be represented as $\mathfrak{R}_m(\mathcal{H}_1) + \mathfrak{R}_m(\mathcal{H}_2)$, where \mathcal{H}_1 and \mathcal{H}_2 are hypothesis space corresponding to the two cascade processes. Because the Rademacher complexity grows

362 exponentially with the depth of the neural network, the network in UE-
 363 LUPI has lower Rademacher complexity (i.e. $\mathfrak{R}_m(\mathcal{H}_1) + \mathfrak{R}_m(\mathcal{H}_2) < \mathfrak{R}_m(\mathcal{H})$).
 364 The reduction of the Rademacher complexity can reduce the Rademacher
 365 generalization bound shown in Equation (6). This indicates the difference
 366 between the true error and training error of the network decreases (i.e. reduce
 367 the generalization error) and thus reduce the model variance. The detailed
 368 theoretical analysis of Proposition 2 is shown in Appendix A. Therefore,
 369 UE-LUPI can improve the performance of the neural network in the strain
 370 reconstruction task.

371 **Implementation in the strain reconstruction task.** The network
 372 architecture in UE-LUPI can be considered as the combination of two cas-
 373 cade subnetworks (N_1 and N_2) corresponding to the RF-to-displacement map
 374 and displacement-to-strain map, respectively. The input of entire network
 375 are two sets of 2D RF data (denoted by R_1 and R_2) collected before and
 376 after compressing the target tissue. Figure 6 shows the network details. The
 377 subnetwork N_1 firstly adopts the separable convolution to obtain the fea-
 378 ture maps of R_1 and R_2 independently and concatenate them. Then, the
 379 high-level contextual information is extracted by locally-connected convolu-
 380 tion (LCC) from this concatenated feature map to represent the difference
 381 between R_1 and R_2 (Chollet, 2017). The feature map is further represented
 382 by dropout, batch normalization (BN), LCC, BN, LCC, dropout, BN, LCC,
 383 BN, LCC, and dropout in sequence. Finally, the 2D feature map are resized
 384 into a vector and then fed into a three-layer fully-connected network (64,
 385 32 and 1 units for three layers). The subnetwork N_2 is composed of the
 386 convolution, BN, convolution, BN, dropout, LCC, BN, LCC and dropout in
 387 sequence. Similar to N_1 , a fully-connected network is finally used to predict
 388 the strain from the vectorized feature produced by the preceding layers.

389 UE-LUPI trains the network only on simulation data and tests it simu-
 390 lation, phantom and real clinical data. Then a part-to-whole training proce-
 391 dure in order to reduce the gradient vanishing influence in the shadow layers
 392 of the network. In this part-to-whole training procedure, UE-LUPI firstly
 393 trains the N_1 and N_2 independently. The loss functions of N_1 (denoted by
 394 L_1) and N_2 (denoted by L_2) is formulated as $L_1 = \sum_{i=1}^{M_1} \sum_{j \in I_1} \|D_{i,j} - D_{i,j}^g\|$
 395 and $L_2 = \sum_{i=1}^{M_2} \sum_{j \in I_2} \|S_{i,j} - S_{i,j}^g\|$, where $D_{i,j}$ and $S_{i,j}$ are the displacement
 396 and strain values at (i, j) estimated by N_1 and N_2 , respectively. $D_{i,j}^g$ and
 397 $S_{i,j}^g$ are corresponding ground truth. M_1 and M_2 are the patch numbers. I_1
 398 and I_2 are the patch regions. Then, UE-LUPI train the entire network by

399 simultaneously using the privileged information and the training labels. The
 400 loss function in Equation (5) can be rewritten as $loss = L_1 + L_2$.

401 3.2. Physically-based data generation for ensuring the causality of the privi- 402 leged information

403 For assuring the causality of the privileged information with the training
 404 examples and labels, UE-LUPI proposes a strategy to generate the triplets of
 405 the RF data, tissue displacement and strain labels by numerical simulation.
 406 This causality comes from the tissue biomechanics and ultrasound physics
 407 used in the numerical simulation which can prompt the data generation pro-
 408 cess to approximate the actual ultrasound elastography process. Moreover,
 409 it can endow the UE-LUPI with conditional independence between the train-
 410 ing examples and labels on the privileged information (see Proposition 1).
 411 In addition, it can address the model overfitting due to data insufficiency, as
 412 well as use all clinical data to validate the model in avoidance of consuming
 413 them in the model training.

414 Specified in the strain reconstruction task, the proposed data generation
 415 aims to generate the synthetic pre- & post-compression RF data with dis-
 416 placement data and strain tags for network training (see Figure 5). In the
 417 numerical simulation, the tissue compression process is the perpendicular
 418 compression of skin by a cylindrical and flat-ended indenter. The soft tissue
 419 is assumed to be homogeneous and isotropic. Based on these assumptions,
 420 we can use the finite element model (mesh the tissue into by massive nodes)
 421 to describe the deformation within the soft tissue under compression, and
 422 compute the spatial locations of the nodes. The change of the node location
 423 in the tissue deformation can provide the displacement field and strain field
 424 that is considered as by the ground truth or the labels of the training data.
 425 Then, these nodes can be regarded as the scatter points in ultrasound sim-
 426 ulation and thus form two kinds of digital phantoms, corresponding to the
 427 tissues before and after compression, respectively. According to Jensen *et al.*
 428 (Jensen and Svendsen, 1992), we can compute the simulated RF signals from
 429 the pre- & post-compression digital phantoms, and then considered these RF
 430 signals as the network input in the training process.

431 4. Experiments and Results

432 In this section, we have shown the effectiveness of our UE-LUPI in simu-
 433 lation, phantom and real clinical data, as well as the data collection, imple-

434 mentation details and evaluation indices.

435 4.1. Data collection

436 4.1.1. Simulation data

437 We produce 200 simulation tissues with $3\text{ cm} \times 5\text{ cm}$ (width \times depth)
 438 according to Section 3.2. In each simulation tissue, 10000 scatter points are
 439 contained with pre-specified locations and an inclusion with random number
 440 of scatter points. The background and inclusions in the simulation tissue has
 441 different values of Young’s module. The background is set 25 kPa, and the
 442 inclusions are set 8 kPa, 14 kPa, 45 kPa and 80 kPa. Then the commercially
 443 available software COMSOL 5.1 is used to calculate the position of all scat-
 444 tering points after compressing the tissue by the finite element method, as
 445 well as the ground truth of tissue displacement field and strain field. Finally,
 446 the 2D RF data of these scattering points before and after tissue compression
 447 is generated by an ultrasound simulation system Field II (Jensen, 1996). The
 448 central frequency and sampling frequency in Field II are set 6 MHz and 40
 449 MHz, respectively.

450 4.1.2. Phantom data

451 We collect the phantom data from a commercially available elasticity
 452 QA phantom (CIRS049, Norfolk, VA, USA). Using a VerasonicsVantage 256
 453 system (Verasonics Inc., WA, USA) equipped with an L12-5 transducer, we
 454 obtain the 2D RF data from 42 regions within this phantom, including four
 455 different inclusions with Young’s module of 8 kPa, 14 kPa, 45 kPa and 80
 456 kPa, as well as a background region with Young’s module of 25 kPa. The
 457 central frequency and sampling frequency of the scanning are set 6.25 MHz
 458 and 40 MHz, respectively.

459 4.1.3. Real clinical data

460 We acquire the four real clinical data (one is for liver and the remaining
 461 are for breast) from a public dataset distributed by Rivaz *et al.* (Rivaz *et al.*,
 462 2011).

463 4.2. Implementation details and evaluation indices

464 All input 2D RF arrays are resized to 2608×128 pixels in both training
 465 and testing stages. The patch sizes in the DRN and SRN are 71×9 and 61×9 ,
 466 respectively. In the training procedure, we use ADAM with the momentum
 467 0.9 as the optimization algorithm. The iteration number is 30 epochs. The

Table 1: Better performance in simulation data of our framework (UE-LUPI) than four state-of-the-art methods.

| | displacement | | strain | | |
|----------------------|-----------------------------------|-------------------------------------|----------------------------------|-----------------------------------|-------------------|
| | SNR _d | RMSE _d (μm) | SNR _e | CNR _e | RMSE _s |
| NCC | 92.69 \pm 8.59 | 3.407 | 10.78 \pm 2.12 | 9.74 \pm 7.82 | 4.202 |
| OF | 98.20 \pm 8.92 | 2.898 | 23.05 \pm 5.51 | 35.83 \pm 7.22 | 1.605 |
| AM-Kalman | 102.75 \pm 7.36 | 2.734 | 19.78 \pm 2.63 | 33.95 \pm 4.75 | 1.893 |
| DL-US | 96.07 \pm 8.86 | 2.815 | 28.68 \pm 6.80 | 62.32 \pm 10.80 | 1.421 |
| Our (UE-LUPI) | 106.57\pm8.35 | 2.502 | 34.76\pm7.45 | 70.45\pm10.45 | 1.251 |

468 learning rate is 0.01. The training dataset contains 100 simulation tissues.
 469 The entire model for training and testing were implemented by TensorFlow
 470 on a NVIDIA GTX1070 GPU.

471 We apply the root-mean-square error (RMSE), as well as two relative in-
 472 dices, elastographic signal-to-noise ratio (SNR_e) and elastographic contrast-
 473 to-noise ratio (CNR_e), to measure the quality of the strain reconstruction
 474 (Srinivasan et al., 2003), because no ground truth can be obtained in phan-
 475 tom data and real clinical data:

$$\begin{aligned} \text{SNR}_e &= 10 \log \frac{e_i}{\sigma_i} \\ \text{CNR}_e &= 10 \log \frac{2(e_b - e_i)^2}{\sigma_b^2 + \sigma_i^2} \end{aligned} \quad (7)$$

476 where e_b and σ_b are the mean value and variance of the background, and e_i
 477 and σ_i the mean value and variance of the inclusion.

478 In order to show that UE-LUPI learns the actual ultrasound elastography
 479 process, we use the RMSE and the displacement signal-to-noise ratio (SNR_d)
 480 to measure the closeness between the output of the intermediate layer in the
 481 network and the generated displacement in simulation data. The SNR_d is
 482 formulated as:

$$\text{SNR}_d = (l_d - l_g)^2 / l_g^2 \quad (8)$$

483 where l_d and l_g are displacement profiles generated by our approach and the
 484 ground truth, respectively (Srinivasan et al., 2003).

485 4.3. Performance of UE-LUPI

486 *Low RMSE and signal-to-noise ratio.* Table 1 shows the high performance
 487 of UE-LUPI in strain reconstruction of simulation tissues, as well as the
 488 displacement estimation. In the strain reconstruction, the low value of RMSE

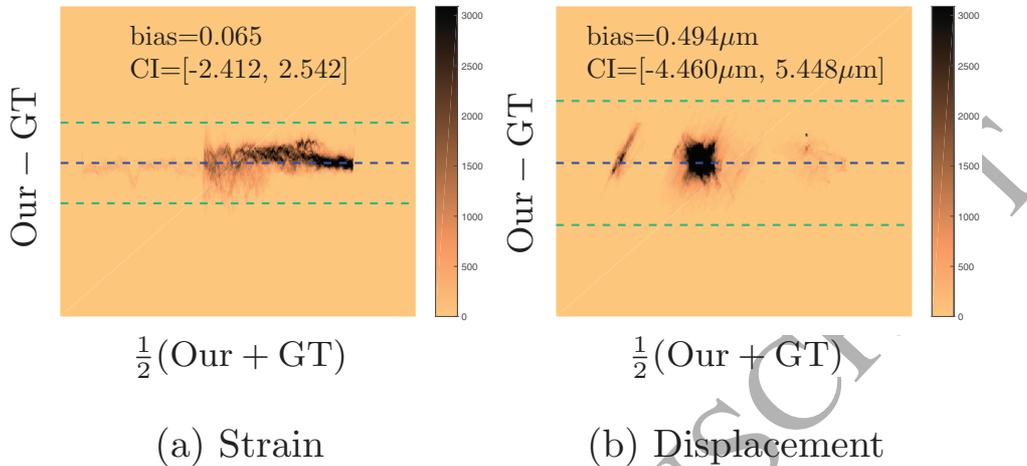


Figure 7: High agreement between our framework (“Our”) and the ground truth (“GT”) evaluated by the Bland-Altman analysis (BA). (a) The results of strain reconstruction. (b) The results of displacement computation. The blue dashed lines are the average values of the bias between the results from our UE-LUPI and the ground truth. The green dashed lines show the 95% confidence interval (CI). The different colors indicate the different density or the frequency of the scatter points in the two-dimensional BA planes. The correspondence between the color and the density are displayed in the color bars. The strain results in BA plots have no unit.

489 (=1.251) indicates that UE-LUPI can reconstruct the strain field with low
 490 error with the ground truth. Moreover, the low values of SNR_e (34.76 ± 7.45
 491 dB) and CNR_e (70.45 ± 10.45 dB) imply that the strain field in the inclusion is
 492 less disturbed by noise, as well as the strain contrast between the background
 493 and the inclusion has a high-level signal-to-noise ratio. In addition, in the
 494 displacement computation, the SNR_d is 106.57 ± 8.35 dB and the RMSE is
 495 0.09. These results show that the error of the displacement field estimated
 496 by UE-LUPI is at a low level.

497 *High agreement.* Figure 7 displays that UE-LUPI is highly agreed with
 498 the ground truth in all simulation data by the Bland-Altman (BA) analysis
 499 (Bland and Altman, 1986). The left and right plots show the results of
 500 tissue strain and displacement, respectively. The different colors indicate the
 501 different density or the frequency of the scatter points in the two-dimensional
 502 BA planes. For the strain reconstruction, the average bias of our results with
 503 respect to the ground truth is 0.065 with 95% confidence interval $[-2.412,$
 504 $2.542]$ in Bland-Altman analysis. For the displacement computation, this

Table 2: Better performance in phantom and patient data of our framework (UE-LUPI) than four state-of-the-art methods.

| | phantom data | | patient data | |
|----------------------|-------------------|--------------------|-------------------|--------------------|
| | SNR _e | CNR _e | SNR _e | CNR _e |
| NCC | 21.34±5.53 | 31.66±23.19 | 16.47±3.45 | 29.44±10.49 |
| OF | 27.53±4.71 | 41.83±9.92 | 21.36±5.53 | 31.66±23.19 |
| AM-Kalman | 29.60±4.02 | 38.21±14.86 | 9.63±12.84 | 36.22±24.98 |
| DL-US | 32.55±5.05 | 45.16±11.23 | 23.26±3.20 | 40.35±18.38 |
| Our (UE-LUPI) | 36.56±3.83 | 48.65±11.43 | 30.44±5.13 | 47.13±20.98 |

505 bias is $0.494 \mu\text{m}$ with 95% confidence interval $[-4.460, 5.448]$.

506 4.4. Outperformance over the state-of-the-art methods

507 The superiority of UE-LUPI has been shown in simulation, phantom and
 508 real clinical data by comparing with four state-of-the-art methods: NCC (Luo
 509 and Konofagou, 2010), OF (Pan et al., 2015), AM-Kalman (Rivaz et al.,
 510 2011) and DL-US (Wu et al., 2018). NCC computed the maximum nor-
 511 malized cross-correlation value between the measurement window and the
 512 best-matched window for estimating the tissue displacement, and then used
 513 the gradient method to obtain the tissue strain. OF used the brightness con-
 514 stancy assumption in the optical flow model to constrain the motion of the
 515 measurement window for estimate tissue displacement, and then utilized the
 516 gradient method to compute the tissue strain. AM-Kalman applied 2D analy-
 517 tic minimization to obtain the tissue displacement from RF data and then
 518 used the Kalman filter to calculate the smooth strain field. DL-US adopted
 519 two neural networks to estimate the displacement field from RF data and the
 520 strain field from the displacement field, respectively. The parameters used
 521 in these state-of-the-art methods are identical with those provided in their
 522 publications.

523 4.4.1. Higher performance in simulation data

524 Table 1 displays that the outperformance of UE-LUPI to the state-of-
 525 the-art methods in the simulation data. The higher SNR and lower RMSE
 526 indicate that UE-LUPI is better than DL-US in the strain reconstruction
 527 and displacement computation (6.08 dB higher for SNR_e, 8.13 dB higher for
 528 CNR_e, and 0.170 lower for the strain RMSE; 10.5 dB higher for SNR_d and
 529 0.313 μm lower for the displacement RMSE). Although in our preliminary
 530 study (Wu et al., 2018), DL-US shows slightly worse than AM-Kalman (6.68

531 dB lower for SNR_d and $0.081 \mu\text{m}$ higher for displacement RMSE), the im-
 532 provement of UE-LUPI with respect to DL-US causes that UE-LUPI shows
 533 better performance than AM-Kalman (3.82 dB higher for SNR_d and 0.232
 534 μm lower for displacement RMSE). UE-LUPI is also superior to AM-Kalman
 535 in the strain reconstruction. Besides, UE-LUPI performs much better than
 536 NCC and OF in both strain reconstruction and displacement computation.
 537 These results show that UE-LUPI is better than these state-of-the-art meth-
 538 ods.

539 4.4.2. Higher performance in phantom and real clinical data

540 Table 2 shows the performance of UE-LUPI is better than the other
 541 methods in phantom data (the increase of $\text{SNR}_e \geq 4.01 \text{ dB}$ and $\text{CNR}_e \geq 3.49$
 542 dB), as well as in real clinical data (the increase of $\text{SNR}_e \geq 7.18 \text{ dB}$ and CNR_e
 543 $\geq 6.78 \text{ dB}$).

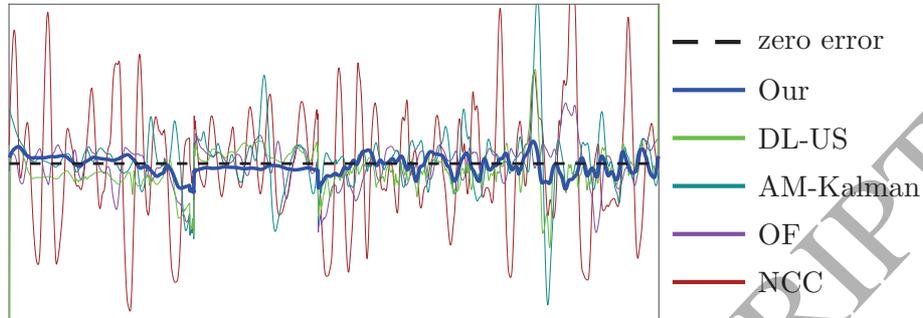
544 4.4.3. Higher performance visualized in the representative data

545 Figure 8 illustrates the higher strain reconstruction performance of UE-
 546 LUPI than the other methods by visualizing a representative simulation data.
 547 The results indicate that UE-LUPI (blue curves) can predict the strain field
 548 closer to the ground truth than NCC (brown), OF (purple), AM-Kalman
 549 (dark cyan), and DL-US (green). Besides, UE-LUPI has the highest lesion-
 550 background contrast and the most homogeneous strain field in all of these
 551 methods.

552 Figure 11 illustrates the higher strain reconstruction performance of UE-
 553 LUPI than the comparative methods by visualizing a representative phantom
 554 data and a representative real clinical data. The strain field reconstructed by
 555 UE-LUPI has better performance on lesion-background contrast and surface
 556 homogeneity than these comparative methods.

557 4.4.4. Higher effectiveness-to-efficiency ratio

558 UE-LUPI has higher effectiveness-to-efficiency ratio (i.e. the ratio of the
 559 evaluation index to the computational time) than all comparative methods.
 560 The effectiveness-to-efficiency ratio of NCC is considered as the baseline value
 561 because NCC, as the first classic ultrasound elastography method (Ophir
 562 et al., 1991), has been widely used in the method comparison in the liter-
 563 ature (e.g. Amidabadi et al. (2018); Hashemi and Rivaz (2017); Shaswary
 564 et al. (2016); Pan et al. (2015)). For simulation data, the average improve-
 565 ment rates of the effectiveness-to-efficiency ratio (the unit is dB/second) with



Error of reconstructed strain profiles with respect to the ground truth

Figure 8: Better performance of our framework (UE-LUPI) than the four comparative methods by displaying the representative results of strain profile. The curve plot displays the errors of the estimated strain profiles and the ground truth in simulation data. It shows that our framework (blue curves) have smaller error to the ground truth (black dashed line) than NCC (brown), OF (purple), AM-Kalman (dark cyan) and DL-US (green). It originates from the lower values of RMSE value and higher values of SNR_d by the UE-LUPI.

566 respect to SNR_e (CNR_e) are 671% (1631%), 506% (1357%), 368% (790%),
 567 58% (172%) in our UE-LUPI, DL-US, AM-Kalman, OF, respectively. For
 568 phantom data, these improvement rates are 308% (266%), 250% (227%),
 569 300% (248%), -10% (-8%). For real clinical data, these improvement rates
 570 are 540% (455%), 347% (333%), 116% (355%), -27% (-39%). These results
 571 can indicate that UE-LUPI can obtain the high-quality strain images when
 572 preserving the high computational efficiency.

573 4.5. Effectiveness of the network configuration in UE-LUPI

574 The performance of different network components in UE-LUPI are shown
 575 by the ablation study (Girshick et al., 2014). Figure 13 shows the results of
 576 the ablation study. Firstly, The results indicate the current configurations
 577 of the proposed network architecture (the pooling strategy, layer number,
 578 dilated rate of the convolution kernel, and patch size) are effective. Then, the
 579 part-to-whole training strategy is replaced by the direct training (i.e. without
 580 training two subnetworks in UE-LUPI independently). The results show that
 581 the SNR_d , SNR_e and CNR_e decrease and the RMSE increase in both strain
 582 reconstruction and displacement computation. Moreover, the LUPI learning
 583 paradigm is removed in UE-LUPI. The results present that SNR_e and CNR_e
 584 largely decrease (9.39 dB and 14.21 dB for simulation data, 6.02 dB and

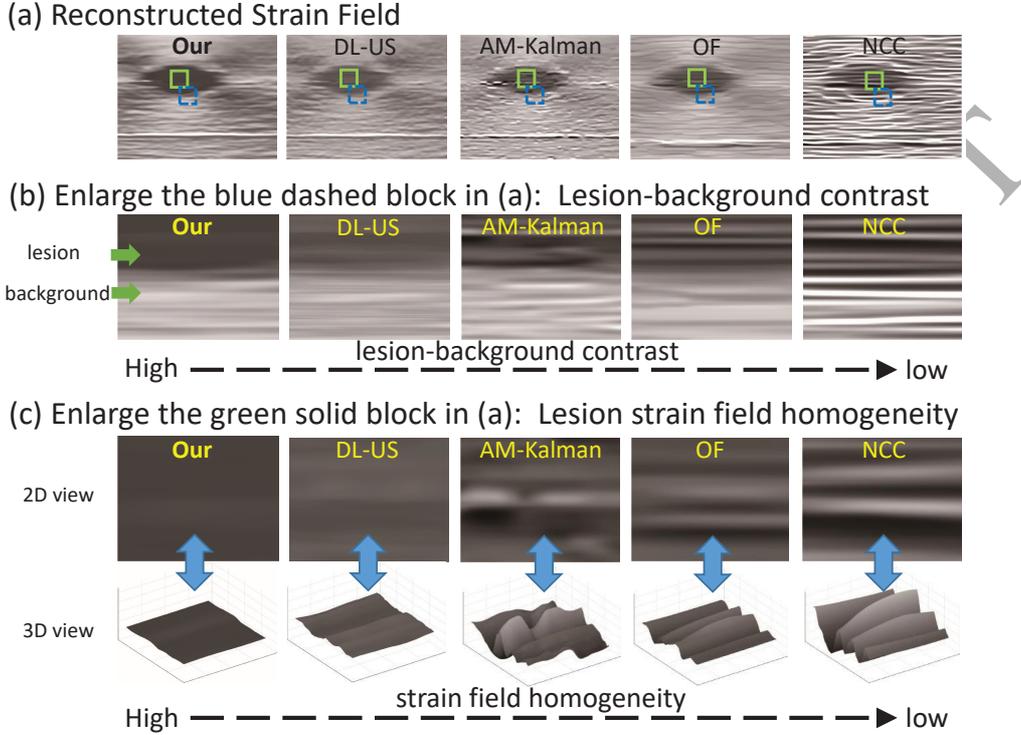


Figure 9: Better performance of our framework (UE-LUPI) than the four comparative methods by displaying the representative results of strain fields in simulation data. (a) shows the completed strain fields. (b) shows the lesion-background contrast in the partial enlargement of blue dashed blocks in (a). (c) shows the image homogeneity and its 3D visualization in the partial enlargement of green blocks in (a). These results present that the strain field estimated by our framework from simulation data has higher lesion-background contrast and more homogeneous appearance than these comparative methods.

585 8.11 dB for phantom data, 11.14 dB and 8.96 dB for real clinical data) when
 586 only preserving the supervision on the output layer of the network. All these
 587 results can demonstrate the effectiveness of the LUPI learning paradigm, the
 588 current network configuration, and the part-to-whole training procedure in
 589 UE-LUPI for strain reconstruction.

590 4.6. Influence from simulation parameters of the physically-based data gen- 591 eration on UE-LUPI

592 The effectiveness of the current configuration in the proposed physically-
 593 based data generation are investigated in four aspects: the scatter amount,

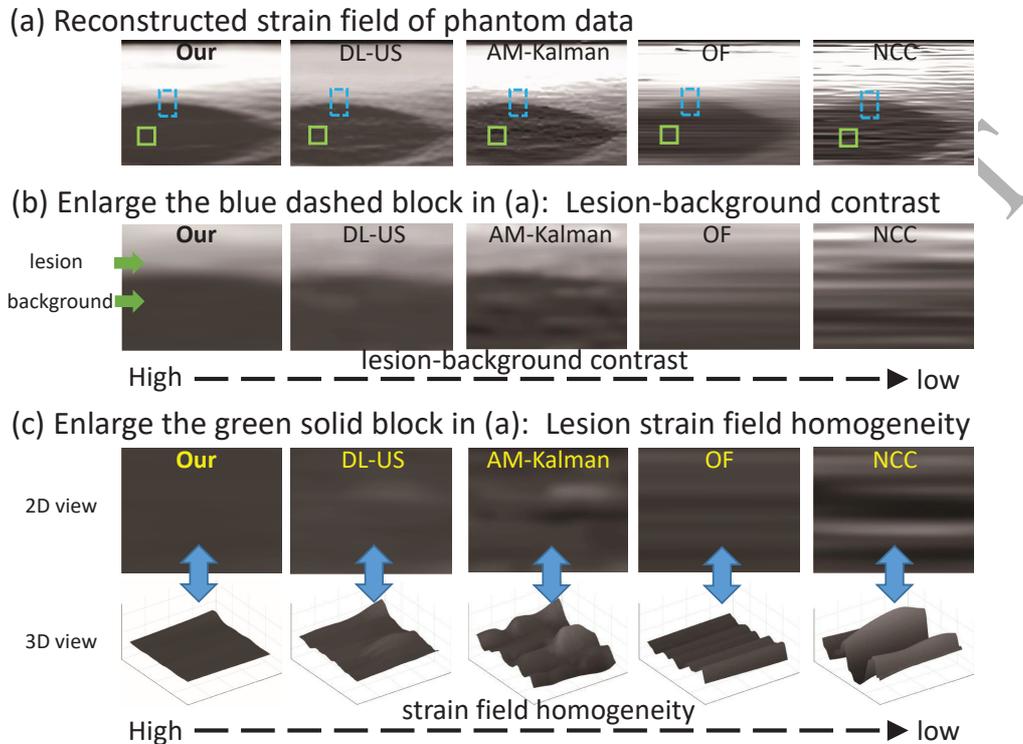


Figure 10: Better performance of our framework (UE-LUPI) than the four comparative methods by displaying the representative results of strain fields in phantom data. (a) shows the completed strain fields. (b) shows the lesion-background contrast in the partial enlargement of blue dashed blocks in (a). (c) shows the image homogeneity and its 3D visualization in the partial enlargement of green blocks in (a). These results present that the strain field estimated by our framework from phantom data has higher lesion-background contrast and more homogeneous appearance than these comparative methods.

594 finite element size, Poisson's ratio and noise corruption.

595 First, the scatter amount is the number of scatter points used to produce
 596 the synthetic tissues in the numerical simulation. Table 3 shows that the
 597 UE-LUPI's performance decreases with the scatter amount (from 1×10^5 to
 598 2×10^4) on simulation data, phantom data and real clinical data, where the
 599 scatter amount used in UE-LUPI is 1×10^5 .

600 Then, the finite element size is the mesh density used in the finite element
 601 model for computing the deformation within the soft tissue under compression.
 602 Table 4 represents that the UE-LUPI's performance decreases when
 603 the finite element size increases (from 0.0005dm to 0.01dm), where the finite

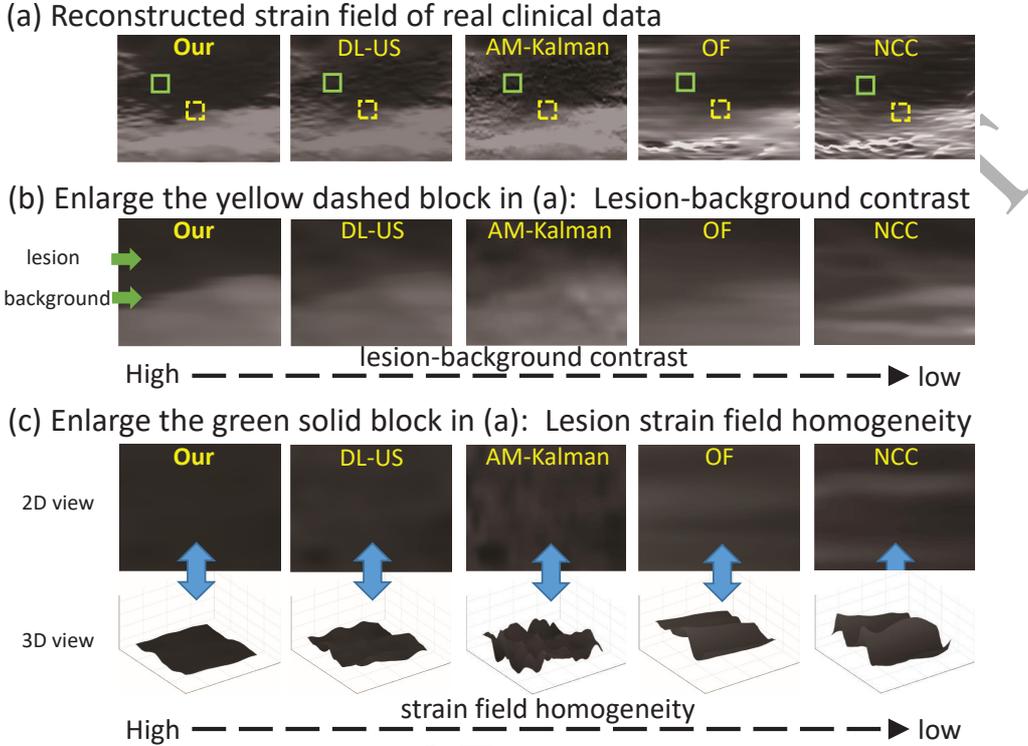


Figure 11: Better performance of our framework (UE-LUPI) than the four comparative methods by displaying the representative results of strain fields in real clinical data. (a) shows the completed strain fields. (b) shows the lesion-background contrast in the partial enlargement of yellow dashed blocks in (a). (c) shows the image homogeneity and its 3D visualization in the partial enlargement of green blocks in (a). These results present that the strain field estimated by our framework from real clinical data has higher lesion-background contrast and more homogeneous appearance than these comparative methods.

604 element size used in UE-LUPI is 0.0005dm.

605 Finally, the Poisson's ratio is the negative of the ratio of transverse strain
606 to axial strain of the tissue, which reflects the material characteristics of
607 the synthetic tissue. Table 5 indicates that UE-LUPI performs well in the
608 Poisson's ratio ranged between 0.4 and 0.5 (i.e. the range of human tissues).

609 4.7. Noise robustness of UE-LUPI

610 The noise influence of UE-LUPI is investigated by adding the extra gaus-
611 sian noise to the RF signals in the training data and then the trained model
612 is tested in simulation, phantom and real clinical data. Four values of the

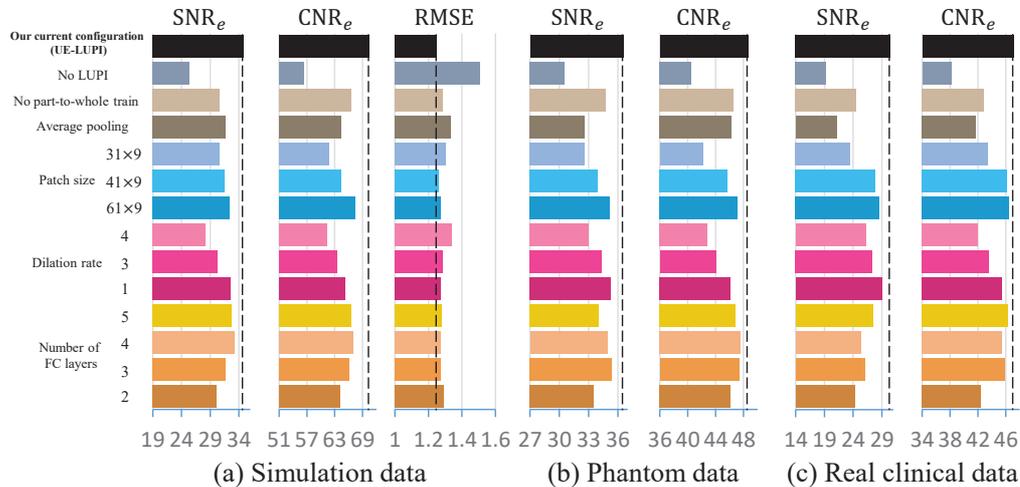


Figure 12: Effectiveness of the UE-LUPI's components in tissue strain reconstruction on (a) simulation data, (b) phantom data and (c) real clinical data. In the ablation study, we evaluate the different configurations of the network architecture, including not using causal privileged information (light gray blue \blacksquare), not using part-to-whole training scheme (light bisque \blacksquare), using average pooling strategy (dark bisque \blacksquare), patch size ($=31\times 9$: light sky blue \blacksquare ; $=41\times 9$: deep sky blue \blacksquare ; $=61\times 9$: dark blue \blacksquare), dilated rate of the convolution kernel ($=4$: pale red \blacksquare ; $=3$: light red \blacksquare ; $=2$: dark red \blacksquare), number of FC layers ($=5$: gold \blacksquare ; $=4$: orange \blacksquare ; $=2$: light tan \blacksquare ; $=1$: dark tan \blacksquare). The better results show the effectiveness of the current configurations of UE-LUPI (shown in black \blacksquare), which applies the causal privileged information, part-to-whole training and max pooling, with dilation rate=1, patch size=71 \times 9 and FC layers=3.

613 noise-to-signal ratio are set: 0.01, 0.02, 0.03 and 0.04. Table 6 shows that
 614 UE-LUPI performs better than all comparative methods even when the noise-
 615 to-signal ratio reaches 0.02. This can demonstrate the noise robustness of
 616 UE-LUPI.

617 5. Discussion

618 5.1. Contributions to clinical community

619 First, UE-LUPI is able to largely reduce the burden to the users in the
 620 tissue elasticity analysis by the quasi-static ultrasound elastography. Because
 621 of the high effectiveness-to-efficiency ratio, UE-LUPI can enable the users to
 622 sidestep the balance between the effectiveness first and the efficiency first,
 623 unlike those must be suffered when using the explicit strain reconstruction

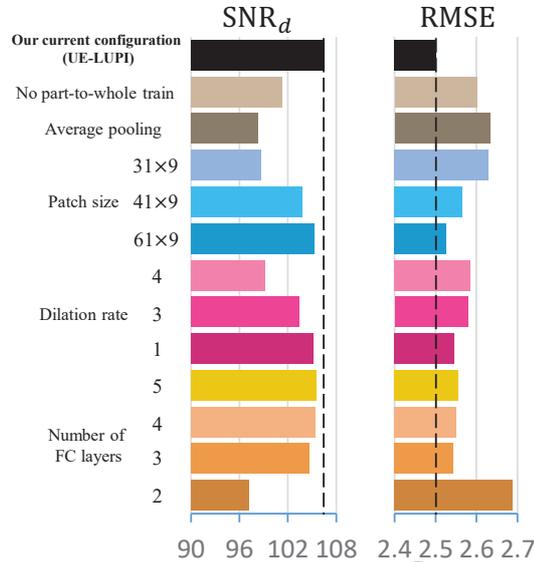


Figure 13: Effectiveness of the UE-LUPI’s components in tissue displacement computation. In the ablation study, we evaluate the different configurations of the network architecture, including not using part-to-whole training scheme (light bisque \square), using average pooling strategy (dark bisque \square), patch size ($=31\times 9$: light sky blue \square ; $=41\times 9$: deep sky blue \square ; $=61\times 9$: dark blue \square), dilated rate of the convolution kernel ($=4$: pale red \square ; $=3$: light red \square ; $=2$: dark red \square), number of FC layers ($=5$: gold \square ; $=4$: orange \square ; $=2$: light tan \square ; $=1$: dark tan \square). The better results show the effectiveness of the current configurations of UE-LUPI (shown in black \blacksquare), which applies the causal privileged information, part-to-whole training and max pooling, with dilation rate=1, patch size= 71×9 , FC layers=3.

624 methods. Contributed to this sidestep, UE-LUPI is able to decrease the ex-
 625 perience requirement for the users about the existing methods. Alternatively,
 626 only the network configuration are needed to be prespecified. Thus, UE-LUPI
 627 can help the inexperienced or fresh users to access the strain reconstruction
 628 task easily and rapidly.

629 Second, UE-LUPI can be readily transferred to other ultrasound elastog-
 630 raphy techniques, such as vibro-acoustography, transient elastography, and
 631 shear wave elasticity imaging. All of these techniques produce the visual-
 632 ization of the tissue elasticity distribution inherently based on the variation
 633 of RF signals with respect to the different biomechanical properties of tis-
 634 sues. Thus, the elasticity reconstruction problem in these methods can be
 635 considered as an inference problem, where the input signals are all RF data

636 and output targets are elasticity-related index. When applying UE-LUPI to
637 other ultrasound elastography methods, the training data and the labels only
638 need to be simply replaced by those in the specified ultrasound elastography
639 method. This potential versatility of UE-LUPI shows its value in clinical
640 ultrasound elastography diagnosis, although UE-LUPI in this work aims at
641 the widespread quasi-static technique (configured in the high-end ultrasound
642 systems of all major manufacturers (Cosgrove et al., 2013)).

643 *5.2. Contributions to medical imaging and machine learning community*

644 We enable the use of the causal LUPI paradigm in network learning by
645 our newly proposed framework for guiding the network to learn the actual
646 target map. The core of this strategy is to provide causal privileged infor-
647 mation to supervise the intermediate layer of the network for correcting the
648 intermediate state of the learning process. By this strategy, it is considered
649 that the learning process of the network can be divided into two simpler
650 cascade subprocesses when the input and output of the learning model are
651 conditionally independent on the its intermediate state. This behavior can
652 narrow the search region of the network parameters owing to the decrease
653 of the Rademacher complexity, and thus induce the map learned by the net-
654 work to evolve towards the actual target map. It further decreases the upper
655 bound of the generalization error and thus reduces the model variance of the
656 network. Specified in the strain reconstruction task, the tissue displacement
657 is a certain intermediate observation in the RF-to-strain process according
658 to tissue biomechanics and ultrasound physics. Hence, using tissue displace-
659 ment as the privileged information can naturally lead the network to learn
660 from the physical causality among RF data, tissue displacement and strain.
661 This strategy can be accessibly applied to any neural network architecture
662 because it is independent on the architecture details.

663 We perform a pioneer study of inferring the physical properties of soft tis-
664 sue from raw physical signals using machine learning. This research direction
665 has not been well established in imaging the elasticity of soft tissue, although
666 a few studies have estimated the elasticity-related physiological indicators
667 from the B-mode ultrasound imaging (Gao et al., 2018a, 2017, 2018b, 2015,
668 2016; Zhao et al., 2018), as well as have investigated the imaging process of
669 tissue by machine learning in MRI and CT (Schlemper et al., 2018, 2017;
670 Chen et al., 2017; Würfl et al., 2016; Xu et al., 2018). In one of embodiments
671 in this research direction, we have developed a framework to infer the tissue
672 strain from pre- & post-compression RF data. Our study inherently aims to

Table 3: Influence on our framework's performance from the training data generated under the different scatter amounts. The current scatter amount is 100000.

| | | Scatter amount (= 100000 in the current configuration) | | | | |
|--------------------|-------------------|---|-------------|-------------|-------------|-------------|
| | | 100000 | 80000 | 60000 | 40000 | 20000 |
| simulation data | SNR _d | 106.57±8.35 | 106.32±7.99 | 106.09±7.07 | 105.31±8.35 | 102.45±7.82 |
| | RMSE _d | 2.502 | 2.513 | 2.561 | 2.892 | 3.019 |
| | SNR _e | 34.76±7.45 | 34.53±7.41 | 33.72±5.56 | 29.86±6.16 | 28.23±6.62 |
| | CNR _e | 70.45±10.45 | 70.34±10.49 | 69.17±9.81 | 57.18±10.98 | 45.04±10.02 |
| | RMSE _s | 1.251 | 1.253 | 1.695 | 2.389 | 2.512 |
| phantom data | SNR _e | 36.56±3.83 | 35.64±3.57 | 34.26±4.83 | 32.02±4.92 | 30.73±4.41 |
| | CNR _e | 48.65±11.43 | 48.28±10.76 | 45.45±11.55 | 39.58±10.45 | 39.05±12.03 |
| patient data | SNR _e | 30.44±5.13 | 29.25±3.48 | 27.69±4.37 | 25.02±5.14 | 23.47±4.98 |
| | CNR _e | 47.13±20.98 | 45.04±18.21 | 41.86±18.84 | 37.34±18.25 | 35.76±16.65 |

Table 4: Influence on our framework's performance from the training data generated under the different finite element sizes.

| | | Finite element size (= 0.0005 in the current configuration) | | | |
|--------------------|-------------------|--|-------------|-------------|-------------|
| | | 0.0005 | 0.001 | 0.005 | 0.01 |
| simulation data | SNR _d | 106.57±8.35 | 106.50±7.77 | 104.14±8.32 | 102.01±8.86 |
| | RMSE _d | 2.502 | 2.570 | 2.994 | 3.187 |
| | SNR _e | 34.76±7.45 | 33.25±8.37 | 31.56±8.18 | 29.97±7.88 |
| | CNR _e | 70.45±10.45 | 66.39±10.27 | 61.64±9.37 | 46.78±9.69 |
| | RMSE _s | 1.251 | 1.547 | 1.994 | 2.271 |
| phantom data | SNR _e | 36.56±3.83 | 33.37±3.02 | 31.6±4.72 | 28.97±5.28 |
| | CNR _e | 48.65±11.43 | 46.94±11.51 | 42.77±12.58 | 41.92±12.58 |
| patient data | SNR _e | 30.44±5.13 | 29.59±5.82 | 27.08±5.01 | 23.28±6.33 |
| | CNR _e | 47.13±20.98 | 45.70±19.96 | 40.85±21.12 | 37.93±21.90 |

Table 5: Influence on our framework’s performance from the training data generated under the different values of Poisson’s ratio.

| | | Poisson’s ratio (= 0.45 in the current configuration) | | | | |
|--------------------|-------------------|--|-------------|-------------|-------------|-------------|
| | | 0.45 | 0.46 | 0.47 | 0.48 | 0.49 |
| simulation data | SNR _d | 106.57±8.35 | 106.56±7.7 | 106.55±7.59 | 106.35±7.97 | 106.10±8.25 |
| | RMSE _d | 2.502 | 2.466 | 2.513 | 2.732 | 2.944 |
| | SNR _e | 34.76±7.45 | 34.62±7.17 | 34.38±7.04 | 33.64±8.35 | 33.94±8.67 |
| | CNR _e | 70.45±10.45 | 70.52±8.99 | 70.36±9.74 | 68.10±10.91 | 66.69±11.33 |
| | RMSE _s | 1.251 | 1.256 | 1.254 | 1.378 | 1.575 |
| phantom data | SNR _e | 36.56±3.83 | 36.26±4.87 | 36.38±3.92 | 35.09±5.65 | 34.47±4.50 |
| | CNR _e | 48.65±11.43 | 47.59±11.14 | 48.23±12.37 | 46.64±10.62 | 45.32±10.79 |
| patient data | SNR _e | 30.44±5.13 | 30.69±3.49 | 29.51±4.61 | 29.02±3.73 | 26.34±3.95 |
| | CNR _e | 47.13±20.98 | 46.27±18.75 | 46.41±19.03 | 47.17±20.57 | 45.42±18.41 |

Table 6: Influence on our framework’s performance from the training data generated under the gaussian noise with different noise-to-signal ratios.

| | | Noise-to-signal ratio (= 0 in the current configuration) | | | | |
|--------------------|-------------------|---|--------------|-------------|-------------|--------------|
| | | 0 | 0.01 | 0.02 | 0.03 | 0.04 |
| simulation data | SNR _d | 106.57±8.35 | 106.55±10.32 | 104.68±8.97 | 103.59±9.36 | 101.95±11.15 |
| | RMSE _d | 2.502 | 2.561 | 2.579 | 2.85 | 3.178 |
| | SNR _e | 34.76±7.45 | 32.09±7.32 | 29.9±7.59 | 27.01±8.02 | 26.32±8.42 |
| | CNR _e | 70.45±10.45 | 68.76±9.48 | 52.77±10.33 | 45.12±11.36 | 39.36±11.62 |
| | RMSE _s | 1.251 | 1.297 | 1.906 | 2.446 | 2.617 |
| phantom data | SNR _e | 36.56±3.83 | 36.44ca±4.02 | 32.62±2.91 | 28.73±2.53 | 27.22±3.04 |
| | CNR _e | 48.65±11.43 | 49.31±10.54 | 45.24±11.04 | 44.07±11.67 | 39.92±12.72 |
| patient data | SNR _e | 30.44±5.13 | 29.27±5.36 | 25.53±4.81 | 24.23±4.11 | 22.75±5.57 |
| | CNR _e | 47.13±20.98 | 47.37±18.66 | 44.65±18.43 | 35.59±15.82 | 31.68±20.12 |

673 solve an attractive machine learning problem applied in a medical imaging
674 scenario: how to infer the biomechanical properties which are difficult to be
675 directly observed from the physical signals.

676 5.3. Reasons for UE-LUPI validity

677 Two goals in the framework design ensure that UE-LUPI has the superi-
678 ority in solving the strain reconstruction task.

679 First, a reasonable learning mechanism should be proposed to learn the
680 relationship among the RF data and tissue strain. To achieve this goal, we
681 are the first to use the deep neural network to extract the strain-related fea-
682 tures from RF data. The choice of the deep neural network is motivated by
683 its effectiveness in the feature learning that has been proved in the literature
684 (Schmidhuber, 2015). Then, we propose a causal privileged information to

685 help the network to learn the actual ultrasound elastography process. This
686 strategy enables to correct the intermediate state in the learning process by
687 formulating this privileged information as the loss term to supervise the in-
688 termediate layer of the network. The causality of the privileged information
689 with the training examples and labels originates from their physical rela-
690 tionship. This can induce the network to gradually learn from the tissue
691 biomechanics and ultrasound physics behind the RF-to-strain map.

692 Second, a reasonable data generation mechanism should be developed to
693 assure the physical causality among the training examples, privileged infor-
694 mation and training labels. This requirement comes from the fact that no
695 direct measurement technique is capable to obtain the training labels (i.e.
696 strain tags) and privileged information (i.e. displacement tags) in strain
697 reconstruction. i.e. all the strain and displacement observations in clin-
698 ical practice are estimated by the existing explicit ultrasound elastography
699 methods. Besides, it is unsuitable to use the results of these existing methods
700 as the privileged information and training data for network learning owing
701 to the incapability to assure the inner physical relationship in ultrasound
702 elastography. To achieve this goal, our data generation strategy is based
703 on tissue-biomechanical and ultrasound-physical simulation. This strategy
704 leads the generation process of the privileged information and training data
705 to approximate the actual ultrasound elastography process and not rely on
706 the existing reconstruction methods. In addition, it can address the data
707 insufficiency problem in the ultrasound elastography in the network training,
708 as well as preserve the rare clinical data for the network testing. Accordingly,
709 the network trained by our data generation strategy is applicable to the real
710 scenario.

711 6. Conclusion

712 In this study, we develop an implicit framework (UE-LUPI) using the
713 deep neural network to reconstruct the strain field in quasi-static ultrasound
714 elastography. The implicitness of UE-LUPI improve the user friendliness,
715 model adaptation, and effectiveness-efficiency ratio. To ensure the suitability
716 of UE-LUPI to the strain reconstruction task, the causal privileged informa-
717 tion is proposed to correct the intermediate state in the learning process for
718 directly guiding the internal feature representation. It can prompt the map
719 learned by the network to evolve towards the actual ultrasound elastography
720 process. Besides, the data generation strategy based on tissue biomechanics

721 and ultrasound physics in UE-LUPI is proposed to approximate the actual
 722 ultrasound elastography process. It can ensure the causality of the privileged
 723 information and address the problem of training data insufficiency. The ex-
 724 perimental results demonstrate the effectiveness of UE-LUPI for the strain
 725 reconstruction task in simulation, phantom and real clinical data.

726 Acknowledgment

727 This work was supported in part by the National Natural Science Founda-
 728 tion of China (61771464, U1801265), Science and Technology Planning
 729 Project of Guangdong Province, China (2018A050506031, 2019B010110001),
 730 and Guangdong Provincial Key Laboratory of Sensor Technology and Biomed-
 731 ical Instrument.

732 Appendix A. Theoretical analysis of Proposition 2

733 Denote x , z and y as the input, intermediate state and output of the
 734 network in UE-LUPI. The standard network aims to learn the map $x \rightarrow y$
 735 (i.e. the function $y = f(x)$), while our network aims to learn the map
 736 $x \rightarrow z \rightarrow y$, (i.e. the function $z = f^{(1)}(x)$ and $y = f^{(2)}(z)$).

737 According to the learning theory (Vapnik, 1998), the true risk $R(h)$ and
 738 the empirical risk $R_n(h)$ of the learning model (e.g. neural network) can be
 739 formulated as

$$\begin{aligned} \text{True risk : } R(h) &= E_x (h(x) - f(x))^2 \\ \text{Empirical risk : } R_m(h) &= \frac{1}{m} \sum_{i=1}^m (h(x_i) - f(x_i))^2 \end{aligned} \quad (\text{A.1})$$

740 where $h(x)$ is a hypothesis in the hypothesis space H , i.e. a function possibly
 741 learned by the network. $f(x)$ is desired function, m is the number of samples.

Then, the generalization error e_g of the learning model can be formulated
 as:

$$e_g = R(\hat{h}) - R_m(\hat{h}) \leq \sup_h |R(h) - R_m(h)| \quad (\text{A.2})$$

742 where $\hat{h} \in H$ is the function learned by the network on the finite sample set.

743 Assume $\|h - f\|_\infty \leq M$ for all $h \in \mathcal{H}$. For any $\delta > 0$ and all $h \in \mathcal{H}$, the
 744 true risk of the learning model has the following Rademacher generalization
 745 bound with probability at least $1 - \delta$ (Bartlett and Mendelson, 2002; Kakade
 746 et al., 2008):

$$R(h) \leq R_m(h) + 4M\mathfrak{R}_m(\mathcal{H}) + M^2\sqrt{\frac{\log \frac{1}{\delta}}{2m}} \quad (\text{A.3})$$

747 The inequality in Equation (A.3) can show the relationship between the true
748 risk and empirical risk for the standard network.

749 For our network with the causal privileged information, denote $h^{(1)}$ and
750 $h^{(2)}$ are the hypothesis corresponding to $f^{(1)}$ and $f^{(2)}$, respectively. The true
751 risk $R'(\hat{h}^{(1)}, \hat{h}^{(2)})$ of our network can be formulated as:

$$\begin{aligned} R'(\hat{h}^{(1)}, \hat{h}^{(2)}) &= E_{\hat{z}, z} (\hat{h}^{(2)}(\hat{z}) - f^{(2)}(z))^2 \\ &= E_{\hat{z}, z} (\hat{h}^{(2)}(\hat{z}) - \hat{h}^{(2)}(z) + \hat{h}^{(2)}(z) - f^{(2)}(z))^2 \\ &\leq E_{\hat{z}, z} (\hat{h}^{(2)}(\hat{z}) - \hat{h}^{(2)}(z))^2 + E_z (\hat{h}^{(2)}(z) - f^{(2)}(z))^2 \\ &\leq E_{\hat{z}, z} (\hat{z} - z)^2 + E_z (\hat{h}^{(2)}(z) - f^{(2)}(z))^2 \\ &= E_x (\hat{h}^{(1)}(x) - f^{(1)}(x))^2 + E_z (\hat{h}^{(2)}(z) - f^{(2)}(z))^2 \\ &\leq \left(R_n^{(1)}(\hat{h}^{(1)}) + 4M\mathfrak{R}_m(\mathcal{H}_1) + M^2\sqrt{\frac{\log \frac{1}{\delta}}{2m}} \right) \\ &\quad + \left(R_n^{(2)}(\hat{h}^{(2)}) + 4M\mathfrak{R}_m(\mathcal{H}_2) + M^2\sqrt{\frac{\log \frac{1}{\delta}}{2m}} \right) \end{aligned} \quad (\text{A.4})$$

752 Thus, this true risk can be rewritten as

$$R'(h) \leq R'_m(h) + 4M(\mathfrak{R}_m(\mathcal{H}_1) + \mathfrak{R}_m(\mathcal{H}_2)) + c \quad (\text{A.5})$$

753 where R'_m is the empirical risk and c is a constant.

754 Owing to the conditional independence between x and y on z , UE-LUPI
755 can divide the learning process of the entire network into two cascade sub-
756 processes without increasing the prediction error. Each subprocess has lower
757 Rademacher complexity (i.e. $\mathfrak{R}_m(\mathcal{H}_1) < \mathfrak{R}_m(\mathcal{H})$ and $\mathfrak{R}_m(\mathcal{H}_2) < \mathfrak{R}_m(\mathcal{H})$)
758 as it corresponds to the network with smaller size. Since the Rademacher
759 complexity grows exponentially with the depth of the neural network at the
760 worst case (Zhang et al., 2017; Sokolic et al., 2016; Golowic et al., 2018;
761 Neyshabur et al., 2015), we have $\mathfrak{R}_m(\mathcal{H}_1) + \mathfrak{R}_m(\mathcal{H}_2) < \mathfrak{R}_m(\mathcal{H})$.

762 Combining it with Equation (A.2), (A.3) and (A.5), we have

$$R'(\hat{h}^{(1)}, \hat{h}^{(2)}) - R'_n(\hat{h}^{(1)}, \hat{h}^{(2)}) < R(h) - R_n(h) \quad (\text{A.6})$$

763 This inequality indicates that UE-LUPI can reduce the generalization error
764 of the network.

765 Reference

766 References

767 Amidabadi, M.G., Ahmad, M.O., Rivaz, H., 2018. Supervised classification
768 of the accuracy of the time delay estimation in ultrasound elastography.
769 IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control
770 65, 21–29.

771 Bamber, J., Cosgrove, D., Dietrich, C.F., Fromageau, J., Bojunga, J., Calli-
772 ada, F., Cantisani, V., Correas, J.M., Onofrio, M.D., Drakonaki, E.E.,
773 Fink, M., Friedrich-Rust, M., Gilja, O.H., Havre, R.F., Jenssen, C.,
774 Klauser, A.S., Ohlinger, R., Saftoiu, A., Schaefer, F., Sporea, I., Piscaglia,
775 F., 2013. EFSUMB guidelines and recommendations on the clinical use of
776 ultrasound elastography. Part 1: Basic principles and technology. Ultra-
777 schall in Medicine 34, 169–184.

778 Barr, R.G., Cosgrove, D., Brock, M., Cantisani, V., Correas, J.M., Postema,
779 A.W., Salomon, G., Tsutsumi, M., Xu, H.X., Dietrich, C.F., 2017.
780 WFUMB guidelines and recommendations for clinical use of ultrasound
781 elastography: Part 5: Prostate. Ultrasound in Medicine and Biology 43,
782 27–48.

783 Barr, R.G., Nakashima, K., Amy, D., Cosgrove, D., Farrokh, A., Schafer, F.,
784 Bamber, J.C., Castera, L., Choi, B.I., Chou, Y.H., Dietrich, C.F., Ding,
785 H., Ferraioli, G., Filice, C., Friedrich-Rust, M., Hall, T.J., Nightingale,
786 K.R., Palmeri, M.L., Shiina, T., Suzuki, S., Sporea, I., Wilson, S., Kudo,
787 M., 2015. WFUMB guidelines and recommendations for clinical use of
788 ultrasound elastography: Part 2: Breast. Ultrasound in Medicine and
789 Biology 41, 1148–1160.

790 Bartlett, P.L., Mendelson, S., 2002. Rademacher and gaussian complexities:
791 Risk bounds and structural results. Journal of Machine Learning Research
792 , 463–482.

793 Bisla, D., Choromanska, A., 2018. Visualbackprop for learning using privi-
794 leged information with CNNs. CoRR abs/1805.09474.

- 795 Bland, J.M., Altman, D.G., 1986. Statistical methods for assessing agreement
796 between two methods of clinical measurement. *The Lancet* 327, 307–310.
- 797 Brusseau, E., Kybic, J., Déprez, J.F., Basset, O., 2008. 2-D locally regu-
798 larized tissue strain estimation from radio-frequency ultrasound images:
799 Theoretical developments and results on experimental data. *IEEE Trans-*
800 *actions on Medical Imaging* 27, 145–160.
- 801 Chang, C.C., Lin, C.H., Lee, C.R., Juan, D.C., Wei, W., Chen, H.T., 2018.
802 Escaping from collapsing modes in a constrained space, in: *European Con-*
803 *ference on Computer Vision (ECCV)*.
- 804 Chen, H., Zhang, Y., Kalra, M.K., Lin, F., Chen, Y., Liao, P., Zhou, J.,
805 Wang, G., 2017. Low-dose CT with a residual encoder-decoder convolu-
806 tional neural network. *IEEE Transactions on Medical Imaging* 36, 2524–
807 2535.
- 808 Chollet, F., 2017. Xception: Deep learning with depthwise separable convo-
809 lutions, in: *IEEE Conference on Computer Vision and Pattern Recognition*
810 *(CVPR)*, pp. 1251–1258.
- 811 Cosgrove, D., Barr, R., Bojunga, J., Cantisani, V., Chammas, M.C., Dighe,
812 M., Vinayak, S., Xu, J.M., Dietrich, C.F., 2017. WFUMB guidelines and
813 recommendations for clinical use of ultrasound elastography: Part 4: Thy-
814 roid. *Ultrasound in Medicine and Biology* 43, 4–26.
- 815 Cosgrove, D., Piscaglia, F., Bamber, J., Bojunga, J., Correas, J.M., Gilja,
816 O.H., Klauser, A.S., Sporea, I., Calliada, F., Cantisani, V., D’Onofrio,
817 M., Drakonaki, E.E., Fink, M., Friedrich-Rust, M., Fromageau, J., Havre,
818 R.F., Jenssen, C., Ohlinger, R., Söftoiu, A., Schaefer, F., Dietrich, C.F.,
819 2013. EFSUMB guidelines and recommendations on the clinical use of
820 ultrasound elastography. Part 2: Clinical applications. *Ultraschall in*
821 *Medicine* 34, 238–253.
- 822 DiBattista, A., Noble, J.A., 2014. An efficient block matching and spectral
823 shift estimation algorithm with applications to ultrasound elastography.
824 *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*
825 61, 407–419.

- 826 Dong, S., Gao, Z., Sun, S., Wang, X., Li, M., Zhang, H., Yang, G., Liu, H.,
827 Li, S., 2018. Holistic and deep feature pyramids for saliency detection, in:
828 British Machine Vision Conference (BMVC).
- 829 Duan, L., Xu, Y., Li, W., Chen, L., Wong, D.W.K., Wong, T.Y., Liu, J.,
830 2014. Incorporating privileged genetic information for fundus image based
831 glaucoma detection, in: International Conference on Medical Image Com-
832 puting and Computer-Assisted Intervention (MICCAI), pp. 204–211.
- 833 Engel, J., Hoffman, M., Roberts, A., 2018. Latent constraints: Learning to
834 generate conditionally from unconditional generative models, in: Interna-
835 tional Conference on Learning Representations (ICLR).
- 836 Fearnhead, P., 2002. Markov chain monte carlo, sufficient statistics, and
837 particle filters. *Journal of Computational and Graphical Statistics* 11,
838 848–862.
- 839 Ferraioli, G., Filice, C., Castera, L., Choi, B.I., Sporea, I., Wilson, S.R.,
840 Cosgrove, D., Dietrich, C.F., Amy, D., Bamber, J.C., Barr, R., Chou,
841 Y.H., Ding, H., Farrokh, A., Friedrich-Rust, M., Hall, T.J., Nakashima,
842 K., Nightingale, K.R., Palmeri, M.L., Schafer, F., Shiina, T., Suzuki, S.,
843 Kudo, M., 2015. WFUMB guidelines and recommendations for clinical
844 use of ultrasound elastography: Part 3: Liver. *Ultrasound in Medicine*
845 *and Biology* 41, 1161–1179.
- 846 Feyereisl, J., Kwak, S., Son, J., Han, B., 2014. Object localization based on
847 structural SVM using privileged information, in: Twenty-seventh Confer-
848 ence on Neural Information Processing Systems (NIPS), pp. 208–216.
- 849 Gao, Z., Li, Y., Sun, Y., Yang, J., Xiong, H., Zhang, H., Liu, X., Wu,
850 W., Liang, D., Li, S., 2018a. Motion tracking of the carotid artery wall
851 from ultrasound image sequences: A nonlinear state-space approach. *IEEE*
852 *Transactions on Medical Imaging* 37, 273–283.
- 853 Gao, Z., Sun, Y., Zhang, H., Ghista, D., Li, Y., Xiong, H., Liu, X., Xie,
854 Y., Wu, W., Li, S., 2016. Carotid artery wall motion estimated from
855 ultrasound imaging sequences using a nonlinear state space approach, in:
856 International Conference on Medical Image Computing and Computer-
857 Assisted Intervention (MICCAI), pp. 98–106.

- 858 Gao, Z., Xiong, H., Liu, X., Zhang, H., Ghista, D., Wu, W., Li, S., 2017.
859 Robust estimation of carotid artery wall motion using the elasticity-based
860 state-space approach. *Medical Image Analysis* 37, 1–21.
- 861 Gao, Z., Xiong, H., Zhang, H., Wu, D., Lu, M., Wu, W., Wong, K.K.L.,
862 Zhang, Y.T., 2015. Motion estimation of common carotid artery wall using
863 a h_∞ filter based block matching method, in: *International Conference*
864 *on Medical Image Computing and Computer-Assisted Intervention (MIC-*
865 *CAI)*, pp. 443–450.
- 866 Gao, Z., Zhang, H., Wang, D., Guo, M., Liu, H., Zhuang, L., Shi, P., 2018b.
867 Robust recovery of myocardial kinematics using dual \mathcal{H}_∞ criteria. *Multi-*
868 *media Tools and Applications* 77, 23043–23071.
- 869 Gatys, L., Ecker, A.S., Bethge, M., 2015. Texture synthesis using convolutional
870 neural networks, in: *Twenty-ninth Conference on Neural Information*
871 *Processing Systems (NIPS)*.
- 872 Gilks, W.R., Richardson, S., Spiegelhalter, D., 1995. *Markov chain Monte*
873 *Carlo in practice*. CRC press.
- 874 Girshick, R., Donahue, J., Darrell, T., Malik, J., 2014. Rich feature hierar-
875 chies for accurate object detection and semantic segmentation, in: *IEEE*
876 *Conference on Computer Vision and Pattern Recognition (CVPR)*, pp.
877 580–587.
- 878 Golowic, N., Rakhlin, A., Shamir, O., 2018. Size-independent sample com-
879 plexity of neural network, in: *Thirty-first Annual Conference On Learning*
880 *Theory (COLT)*.
- 881 Goudet, O., Kalainathan, D., Caillou, P., Guyon, I., Lopez-Paz, D., Sebag,
882 M., 2018. Learning functional causal models with generative neural net-
883 works, in: Escalante, H.J., Escalera, S., Guyon, I., Baró, X., Güçlütürk, Y.,
884 Güçlü, U., van Gerven, M. (Eds.), *Explainable and Interpretable Models*
885 *in Computer Vision and Machine Learning*. Springer, pp. 39–80.
- 886 Guo, Y., Liu, Y., Oerlemans, A., Lao, S., Wu, S., Lew, M.S., 2016. Deep
887 learning for visual understanding: A review. *Neurocomputing* , 27–48.
- 888 Hasan, M.K., Anas, E.M.A., Alam, S.K., Lee, S.Y., 2012. Direct mean strain
889 estimation for elastography using nearest-neighbor weighted least-squares

- 890 approach in the frequency domain. *Ultrasound in Medicine and Biology*
891 38, 1759–1777.
- 892 Hashemi, H.S., Rivaz, H., 2017. Global time-delay estimation in ultrasound
893 elastography. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Fre-*
894 *quency Control* 64, 1625–1636.
- 895 Hjelm, R.D., Fedorov, A., Lavoie-Marchildon, S., Grewal, K., Bachman, P.,
896 Trischler, A., Bengio, Y., 2019. Learning deep representations by mutual
897 information estimation and maximization, in: *International Conference on*
898 *Learning Representations (ICLR)*.
- 899 Hoerig, C., Ghaboussi, J., Insana, M.F., 2017. An information-based machine
900 learning approach to elasticity imaging. *Biomechanics and Modeling in*
901 *Mechanobiology* 16, 805–822.
- 902 Jensen, J.A., 1996. FIELD: A program for simulating ultrasound systems,
903 in: *10th Nordicbaltic Conference on Biomedical Imaging*, pp. 351–353.
- 904 Jensen, J.A., Svendsen, N.B., 1992. Calculation of pressure fields from arbi-
905 trarily shaped, apodized, and excited ultrasound transducers. *IEEE Trans-*
906 *actions on Ultrasonics, Ferroelectrics, and Frequency Control* 39, 262–267.
- 907 Jiang, J., Hall, T.J., 2015. A coupled subsample displacement estimation
908 method for ultrasound-based strain elastography. *Physics in Medicine and*
909 *Biology* 60, 8347–8364.
- 910 Kakade, S.M., Sridharan, K., Tewari, A., 2008. On the complexity of linear
911 prediction: Risk bounds, margin bounds, and regularization, in: *Twenty-*
912 *second Conference on Neural Information Processing Systems (NIPS)*.
- 913 Kibria, G., Riyaz, H., 2018. GLUENet: Ultrasound elastography using convo-
914 lutional neural network, in: Stoyanov, D., Taylor, Z., Aylward, S., Tavares,
915 J.M.R., Xiao, Y., Simpson, A., Martel, A., Maier-Hein, L., Li, S., Rivaz,
916 H., Reinertsen, I., Chabanas, M., Farahani, K. (Eds.), *Simulation, Image*
917 *Processing, and Ultrasound Systems for Assisted Diagnosis and Naviga-*
918 *tion*, Springer. pp. 21–28.
- 919 Koller, D., Friedman, N., 2009. *Probabilistic Graphical Models: Principles*
920 *and Techniques*. The MIT Press.

- 921 Kuzmin, A., Zakrzewski, A.M., Anthony, B.W., Lempitsky, V., 2015. Multi-
922 frame elastography using a handheld force-controlled ultrasound probe.
923 IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control
924 62, 1486–1500.
- 925 Lambert, J., Sener, O., Savarese, S., 2018. Deep learning under privileged
926 information using heteroscedastic dropout, in: IEEE Conference on Com-
927 puter Vision and Pattern Recognition (CVPR), pp. 8886–8895.
- 928 Lapin, M., Hein, M., Schiele, B., 2014. Learning using privileged information:
929 SVM+ and weighted SVM. Neural Networks , 95–108.
- 930 Lee, C.Y., Xie, S., Gallagher, P.W., Zhang, Z., Tu, Z., 2015. Deeply-
931 supervised nets, in: International Conference on Artificial Intelligence and
932 Statistics (AISTATS), pp. 562–570.
- 933 Lu, M., Wu, D., Lin, W., Li, W., Zhang, H., Huang, W., 2014. A stochastic
934 filtering approach to recover strain images from quasi-static ultrasound
935 elastography. BioMedical Engineering OnLine 13, 1–17.
- 936 Luo, J., Konofagou, E.E., 2010. A fast normalized cross-correlation calcu-
937 lation method for motion estimation. IEEE Transactions on Ultrasonics,
938 Ferroelectrics, and Frequency Control 57, 1347–1357.
- 939 Meng, F., Shi, J., Gong, B., Zhang, Q., Guo, L., Wang, D., Xu, H., 2018.
940 B-mode ultrasound based diagnosis of liver cancer with CEUS images as
941 privileged information, in: 40th Annual International Conference of the
942 IEEE Engineering in Medicine and Biology Society (EMBC), pp. 3124–
943 3127.
- 944 Neyshabur, B., Tomioka, R., Srebro, N., 2015. Norm-based capacity control
945 in neural networks, in: Twenty-eighth Annual Conference On Learning
946 Theory (COLT).
- 947 Omidyeganeh, M., Xiao, Y., Ahmad, M.O., Rivaz, H., 2017. Estimation
948 of strain elastography from ultrasound radio-frequency data by utilizing
949 analytic gradient of the similarity metric. IEEE Transactions on Medical
950 Imaging 36, 1347–1358.

- 951 Ophir, J., Céspedes, I., Ponnekanti, H., Yazdi, Y., Li, X., 1991. Elastogra-
952 phy: A quantitative method for imaging the elasticity of biological tissues.
953 *Ultrasonic Imaging* 13, 111–134.
- 954 Pan, X., Liu, K., Shao, J., Gao, J., Huang, L., Bai, J., Luo, J., 2015. Per-
955 formance comparison of rigid and affine models for motion estimation us-
956 ing ultrasound radio-frequency signals. *IEEE Transactions on Ultrasonics,*
957 *Ferroelectrics, and Frequency Control* 62, 1928–1943.
- 958 Pechyony, D., Vapnik, V., 2010. On the theory of learning with privileged
959 information, in: *Twenty-fourth Conference on Neural Information Pro-*
960 *cessing Systems (NIPS)*.
- 961 Pellet-Barakat, C., Frouin, F., Insana, M.F., Herment, A., 2004. Ultrasound
962 elastography based on multiscale estimations of regularized displacement
963 fields. *IEEE Transactions on Medical Imaging* 23, 153–163.
- 964 Rivaz, H., Boctor, E.M., Choti, M.A., Hager, G.D., 2011. Real-time regular-
965 ized ultrasound elastography. *IEEE Transactions on Medical Imaging* 30,
966 928–945.
- 967 Rivaz, H., Boctor, E.M., Choti, M.A., Hager, G.D., 2014. Ultrasound elas-
968 tography using multiple images. *Medical Image Analysis* 18, 314–329.
- 969 Rivaz, H., Foroughi, P., Fleming, I., Zellars, R., Boctor, E., Hager, G., 2009.
970 Tracked regularized ultrasound elastography for targeting breast radio-
971 therapy, in: *International Conference on Medical Image Computing and*
972 *Computer-Assisted Intervention (MICCAI)*, pp. 507–515.
- 973 Schlemper, J., Caballero, J., Hajnal, J.V., Price, A., Rueckert, D., 2017. A
974 deep cascade of convolutional neural networks for mr image reconstruction,
975 in: *International Conference on Information Processing in Medical Imaging*
976 *(IPMI)*, pp. 647–658.
- 977 Schlemper, J., Caballero, J., Hajnal, J.V., Price, A.N., Rueckert, D., 2018.
978 A deep cascade of convolutional neural networks for dynamic MR image
979 reconstruction. *IEEE Transactions on Medical Imaging* 37, 491–503.
- 980 Schmidhuber, J., 2015. Deep learning in neural networks: An overview.
981 *Neural Networks* 62, 85–117.

- 982 Sharmanska, V., Quadrianto, N., Lampert, C.H., 2013. Learning to rank
983 using privileged information, in: IEEE International Conference on Com-
984 puter Vision (ICCV), pp. 825–832.
- 985 Sharmanska, V., Quadrianto, N., Lampert, C.H., 2014. Learning to transfer
986 privileged information. CoRR abs/1410.0389.
- 987 Shaswary, E., Xu, Y., Tavakkoli, J., 2016. Performance study of a new
988 time-delay estimation algorithm in ultrasonic echo signals and ultrasound
989 elastography. *Ultrasonics* 69, 11–18.
- 990 Shi, J., Xue, Z., Dai, Y., Peng, B., Dong, Y., Zhang, Q., Zhang, Y., 2018.
991 Cascaded multi-column RVFL+ classifier for single-modal neuroimaging-
992 based diagnosis of parkinson’s disease. *IEEE Transactions on Biomedical*
993 *Engineering* (in press) .
- 994 Shi, Z., Kim, T.K., 2017. Learning and refining of privileged information-
995 based RNNs for action recognition from depth sequences, in: IEEE Con-
996 ference on Computer Vision and Pattern Recognition (CVPR), pp. 4684–
997 4693.
- 998 Shiina, T., Nightingale, K.R., Palmeri, M.L., Hall, T.J., Bamber, J.C., Barr,
999 R.G., Castera, L., Choi, B.I., Chou, Y.H., Cosgrove, D., Dietrich, C.F.,
1000 Ding, H., Amy, D., Farrokh, A., Ferraioli, G., Filice, C., Friedrich-Rust,
1001 M., Nakashima, K., Schafer, F., Sporea, I., Suzuki, S., Wilson, S., Kudo,
1002 M., 2015. WFUMB guidelines and recommendations for clinical use of
1003 ultrasound elastography: Part 1: Basic principles and terminology. *Ultra-*
1004 *sound in Medicine and Biology* 41, 1126–1147.
- 1005 Sokolic, J., Giryes, R., Sapiro, G., Rodrigues, M.R.D., 2016. Lessons from
1006 the rademacher complexity for deep learning, in: International Conference
1007 on Learning Representations Workshop.
- 1008 Srinivasan, S., Righetti, R., Ophir, J., 2003. Trade-offs between the axial
1009 resolution and the signal-to-noise ratio in elastography. *Ultrasound in*
1010 *Medicine and Biology* 29, 847–866.
- 1011 Techavipoo, U., Varghese, T., 2004. Wavelet denoising of displacement esti-
1012 mates in elastography. *Ultrasound in Medicine and Biology* 30, 477–491.

- 1013 Thung, K.H., Wee, C.Y., 2018. A brief review on multi-task learning. Mul-
1014 timedia Tools and Applications 77, 29705–29725.
- 1015 Treece, G., Lindop, J., Chen, L., Housden, J., Prager, R., Gee, A., 2011.
1016 Real-time quasi-static ultrasound elastography. Interface Focus 1, 540–
1017 552.
- 1018 Vapnik, V., Izmailov, R., 2015. Learning using privileged information: Sim-
1019 ilarity control and knowledge transfer. Journal of Machine Learning Re-
1020 search , 2023–2049.
- 1021 Vapnik, V., Vashist, A., 2009. A new learning paradigm: Learning using
1022 privileged information. Neural Networks , 544–557.
- 1023 Vapnik, V.N., 1998. Statistical Learning Theory. Wiley-Interscience.
- 1024 Wainwright, M.J., Jordan, M.I., 2008. Graphical models, exponential fami-
1025 lies, and variational inference. Foundations and Trends in Machine Learn-
1026 ing 1, 1–305.
- 1027 Walker, W.F., Trahey, G.E., 1995. A fundamental limit on delay estimation
1028 using partially correlated speckle signals. IEEE Transactions on Ultrason-
1029 ics, Ferroelectrics, and Frequency Control 42, 301–308.
- 1030 Wu, S., Gao, Z., Luo, J., Liu, Z., Zhang, H., Li, S., 2018. Direct reconstruc-
1031 tion of ultrasound elastography using an end-to-end deep neural network,
1032 in: International Conference on Medical Image Computing and Computer-
1033 Assisted Intervention (MICCAI), pp. 374–382.
- 1034 Würfl, T., Ghesu, F.C., Christlein, V., Maier, A., 2016. Deep learning com-
1035 puted tomography, in: International Conference on Medical Image Com-
1036 puting and Computer-Assisted Intervention (MICCAI), pp. 432–440.
- 1037 Xu, C., Xu, L., Gao, Z., Zhao, S., Zhang, H., Zhang, Y., Du, X., Zhao, S.,
1038 Ghista, D., Liu, H., Li, S., 2018. Direct delineation of myocardial infarction
1039 without contrast agents using a joint motion feature learning architecture.
1040 Medical Image Analysis 50, 82–94.
- 1041 Yuan, J., Zhang, H., Lu, M., Chen, S., Liu, H., 2011. A strain-based ul-
1042 trasound elastography using phase shift with prior estimates and meshfree
1043 shape function, in: IEEE International Symposium on Biomedical Imag-
1044 ing: From Nano to Macro (ISBI), pp. 532–535.

- 1045 Yuan, L., Pedersen, P.C., 2015. Analytical phase-tracking-based strain esti-
1046 mation for ultrasound elasticity. *IEEE Transactions on Ultrasonics, Fer-*
1047 *roelectrics, and Frequency Control* 62, 185–207.
- 1048 Zahiri-Azar, R., Salcudean, S.E., 2006. Motion estimation in ultrasound
1049 images using time domain cross correlation with prior estimates. *IEEE*
1050 *Transactions on Biomedical Engineering* 53, 1990–2000.
- 1051 Zakaria, T., Qin, Z., Maurice, R.L., 2010. Optical-flow-based b-mode elas-
1052 tography: Application in the hypertensive rat carotid. *IEEE Transactions*
1053 *on Medical Imaging* 29, 570–578.
- 1054 Zaleska-Dorobisz, U., Kaczorowski, K., Pawluś, A., Puchalska, A., Ingot,
1055 M., 2014. Ultrasound elastography: review of techniques and its clinical
1056 applications. *Advances in Clinical and Experimental Medicine* 23, 645–655.
- 1057 Zhang, C., Bengio, S., Hardt, M., Recht, B., Vinyals, O., 2017. Under-
1058 standing deep learning requires rethinking generalization, in: *International*
1059 *Conference on Learning Representations (ICLR)*.
- 1060 Zhang, K., Wang, Z., Zhang, J., Schölkopf, B., 2016. On estimation of
1061 functional causal models: General results and application to the post-
1062 nonlinear causal model. *ACM Transactions on Intelligent Systems and*
1063 *Technology* 7, 1–23.
- 1064 Zhang, Y., Lv, Q., Yin, Y., Xie, M., Xiang, F., Lu, C., Yan, T., Li, W., Xu,
1065 H., Huang, Y., 2009. The value of ultrasound elastography in differential
1066 diagnosis of superficial lymph nodes. *Frontiers of Medicine in China* 3,
1067 368–374.
- 1068 Zhang, Y., Yang, Q., 2018. A survey on multi-task learning. *CoRR*
1069 [abs/1707.08114](https://arxiv.org/abs/1707.08114).
- 1070 Zhao, S., Gao, Z., Zhang, H., Xie, Y., Luo, J., Ghista, D., Wei, Z., Bi,
1071 X., Xiong, H., Xu, C., Li, S., 2018. Robust segmentation of intima-media
1072 borders with different morphologies and dynamics during the cardiac cycle.
1073 *IEEE Journal of Biomedical and Health Informatics* 22, 1571–1582.
- 1074 Zheng, X., Shi, J., Ying, S., Zhang, Q., Li, Y., 2016. Improving single-
1075 modal neuroimaging based diagnosis of brain disorders via boosted privi-
1076 leged information learning framework, in: *MICCAI Workshop on Machine*
1077 *Learning in Medical Imaging (MLMI)*, pp. 95–103.

1078 Zhu, Y., Hall, T.J., 2002. A modified block matching method for real-time
1079 freehand strain imaging. *Ultrasonic Imaging* 24, 161–176.

ACCEPTED MANUSCRIPT