Learning from Noisy Label Statistics: Detecting High Grade Prostate Cancer in Ultrasound Guided Biopsy

Shekoofeh Azizi[1(✉)], Pingkun Yan[2], Amir Tahmasebi[3], Peter Pinto[4], Bradford Wood[4], Jin Tae Kwak[5], Sheng Xu[4], Baris Turkbey[4], Peter Choyke[4], Parvin Mousavi[6], and Purang Abolmaesumi[1]

[1] University of British Columbia, Vancouver, Canada
shazizi@ece.ubc.ca
[2] Rensselaer Polytechnic Institute, Troy, USA
[3] Philips Research North America, Cambridge, USA
[4] National Institutes of Health, Bethesda, USA
[5] Sejong University, Seoul, Korea
[6] Queen's University, Kingston, Canada

Abstract. The ubiquity of noise is an important issue for building computer-aided diagnosis models for prostate cancer biopsy guidance where the histopathology data is sparse and not finely annotated. We propose a solution to alleviate this challenge as a part of Temporal Enhanced Ultrasound (TeUS)-based prostate cancer biopsy guidance method. Specifically, we embed the prior knowledge from the histopathology as the soft labels in a two-stage model, to leverage the problem of diverse label noise in the ground-truth. We then use this information to accurately detect the grade of cancer and also to estimate the length of cancer in the target. Additionally, we create a Bayesian probabilistic version of our network, which allows evaluation of model uncertainty that can lead to any possible misguidance during the biopsy procedure. In an *in vivo* study with 155 patients, we analyze data from 250 suspicious cancer foci obtained during fusion biopsy. We achieve the average area under the curve of 0.84 for cancer grading and mean squared error of 0.12 in the estimation of tumor in biopsy core length.

Keywords: Temporal enhanced ultrasound · Prostate cancer
Recurrent neural networks

1 Introduction

The ultimate diagnosis for prostate cancer is through histopathology analysis of prostate biopsy, guided by either Transrectal Ultrasound (TRUS), or fusion of TRUS with multi-parametric Magnetic Resonance Imaging (mp-MRI) [14,15]. Computer-aided diagnosis models for detection of prostate cancer and guidance of biopsy involve both ultrasound (US)- and mp-MRI-based tissue

A. F. Frangi et al. (Eds.): MICCAI 2018, LNCS 11073, pp. 21–29, 2018.
https://doi.org/10.1007/978-3-030-00937-3_3

characterization. mp-MRI has high sensitivity in detection of prostate lesions but low specificity [1,10], hence, limiting its utility in detecting disease progression over time [15]. US-based tissue characterization methods focus on the analysis of texture [11] and spectral features [7] within a single ultrasound frame, Doppler imaging and elastography [13]. Temporal Enhanced Ultrasound (TeUS), involving a time-series of ultrasound RF frames captured from insonification of tissue over time [6], has enabled the depiction of patient-specific cancer likelihood maps [2,3,5,12]. Despite promising results in detecting prostate cancer, accurate characterization of aggressive lesions from indolent ones is an open problem and requires refinement.

The goodness of models built based on all the above analyses depends on detailed, noise-free annotations of ground-truth labels from pathology. However, there are two key challenges with the ground-truth. First, histopathology data used for training of the models is sparsely annotated with the inevitable ubiquity of noise. Second, the heterogeneity in morphology and pathology of the prostate itself contributes as a source of inaccuracy in labeling.

In this paper, we propose a method to address the challenge of sparse and noisy histopathology ground-truth labels to improve TeUS-based prostate biopsy guidance. The contributions of the paper are: (1) Employing prior histopathology knowledge to estimate ground-truth probability vectors as soft labels. We then use these soft labels as a replacement for the sparse and noisy labels for training a two-stage Recurrent Neural Networks (RRN)-based model; (2) Using the new ground-truth probability vectors to accurately estimate the tumor in biopsy core length; and (3) A strategy for the depiction of new patient-specific colormaps for biopsy guidance using the estimated model uncertainty.

2 Materials

Data Acquisition and Preprocessing. We use TeUS data from 250 biopsy targets of 155 subjects. All subjects were identified as suspicious for cancer in a preoperative mp-MRI examination. The subjects underwent MRI-guided ultrasound biopsy using UroNav (Invivo Corp., FL) MR-US fusion system [14]. Prior to biopsy sampling from each target, the ultrasound transducer is held steady for 5 seconds to obtain $T = 100$ frames of TeUS RF data. This procedure is followed by firing the biopsy gun to acquire a tissue specimen. Histopathology information of each biopsy core is used as the gold-standard for generating a label for that core. For each biopsy target, we analyze an area of $2\,\text{mm} \times 10\,\text{mm}$ around the target, along with the projected needle path. We divide this region into 80 equally-sized Regions of Interest (ROIs) of $0.5\,\text{mm} \times 0.5\,\text{mm}$. For each ROI, we generate a sequence of TeUS data, $\mathbf{x}^{(i)} = (x_1^{(i)}, ..., x_T^{(i)}), T = 100$ by averaging over all the time series values within a given ROI of an ultrasound frame (Fig. 2). An individual TeUS sequence is constituted of echo-intensity values $x_t^{(i)}$ for each time step, t. We also augment the training data ($\mathcal{D}_{train}$) by generating ROIs using a sliding window of size $0.5\,\text{mm} \times 0.5\,\text{mm}$ over the target region with the step size of 0.1 mm, which results in 1,536 ROIs per target.

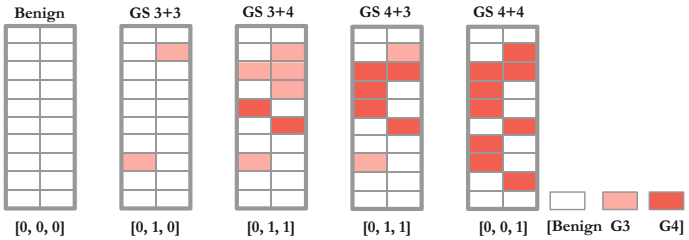


Fig. 1. Illustration of noisy and not finely annotated ground-truth label. The exact location of the cancerous ROI in the core, the ratio of the different Gleason grade, and the exact location of the Gleason grades are unknown and noisy. The bottom vectors show one of the possible multi-label binarization approaches.

Ground-Truth Labeling. Histopathology reports include the length of cancer in the biopsy core and a Gleason Score (GS) [13]. The GS is reported as a summation of the Gleason grades of the two most common cancer patterns in the tissue specimen. Gleason grades range from 1 (normal) to 5 (aggressive cancerous). The histopathology reports a measure of the statistical distribution of cancer in the cancer foci. The ground-truth is noisy and not finely annotated to show the exact location of the cancerous tissue in the core (Fig. 1). Therefore, the exact grade of each ROI in a core is not available while the overarching goal is to determine the ROI-level grade of the specimen. In our dataset, 78 biopsy cores are cancerous with GS $3+3$ or higher, where 26 of those are labeled as clinically significant cancer with GS $\geq 4+3$. The remaining 172 cores are benign.

3 Method

3.1 Discriminative Model

Let $\mathcal{D} = \{(\mathbf{x}^{(i)}, \mathbf{y}^{(i)})\}_{i=1}^{|\mathcal{D}|}$ represent the collection of all ROIs, where $\mathbf{x}^{(i)}$ is the i^{th} TeUS sequence with length T and is labeled as $\mathbf{y}^{(i)}$) corresponding to a cancer grade. The objective is to develop a probabilistic model to discriminate between cancer grades using noisy and not well-annotated data in $\mathcal{D}$. For this purpose, we propose a two-stage approach to consider the diverse nature of noise in the ground-truth labeling: benign vs. all grades of cancer and the mixture of cancer grades. The goal of the first stage is to mine the data points with non-cancerous tissue in the presence of possible noise where several theoretical studies have shown the robustness of the binary classification accuracy to the simple and symmetric label noise [8]. The goal of the second stage is to learn from the noisy label statistic in cancerous cores by suppressing the influence of noise using a soft label. In the core of the approach, we use deeply connected RNN layers to explicitly model the temporal information in TeUS followed by a fully connected layer to map the sequence to a posterior over classes. Each RNN layer includes $T = 100$ homogeneous hidden units (*i.e.*, traditional/vanilla RNN, LSTM or GRU cells) to capture temporal changes in data. Given the

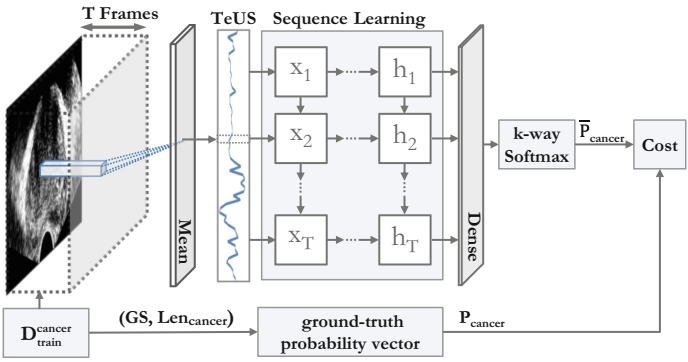


Fig. 2. Overview of the second stage of the method: the goal of is to assign a pathological score to a sample. To mitigate the problem of noisy labels, we embed the length of the cancer in the ground-truth probability vector as a soft label.

input sequence $\mathbf{x} = (x_1, , ..., x_T)$, RNN computes the hidden vector sequence $\mathbf{h} = (h_1, , ..., h_T)$ in the sequence learning step. This hidden vector, $\mathbf{h}$ is a function of the input sequence $\mathbf{x}$, model parameters, Θ, and time.

Stage 1: Detection of Benign Samples: Let $y_b^{(i)} \in \{0, 1\}$ indicate the corresponding binary label for $\mathbf{x}^{(i)}$, where zero and one indicate benign and cancer outcome, respectively. We aim to learn a mapping from $\mathbf{x}^{(i)}$ to $y_b^{(i)}$ in a supervised manner. After the sequence learning step, the final node generates the posterior probability for the given sequence:

$$\overline{y}_b^{(i)} = \arg\max_j \mathcal{S}(\mathbf{z}_j^{(i)}),\ j \in \{0, 1\},\ \mathbf{z}^{(i)} = \mathbf{W}_s^T \mathbf{h} + \mathbf{b}_s, \tag{1}$$

where $\mathbf{W}_s$ and $\mathbf{b}_s$ are the weight and bias of the fully-connected layer and $\mathcal{S}$ is the softmax function, which in our binary classification case is equivalent to the logistic function, and $\overline{y}_b^{(i)}$ indicates the predicted label. The optimization criterion for the network is to minimize the binary cross-entropy between $y_b^{(i)}$ and $\overline{y}_b^{(i)}$ over all training samples.

Stage 2: Grading of Cancerous Samples: The goal of this stage is to assign a pathological score to $\mathcal{D}_{train}^{cancer} = \{(\mathbf{x}^{(i)}, y_b^{(i)} \in \mathcal{D}_{train} \mid y_b^{(i)} = 1\}_{i=1}^N$. Here, unlike the first stage, we are facing a multi-label classification task with sparse labeling (Fig. 1). The histopathology reports include two informative parts: (1) Gleason score which implies any of the possible labels $\Omega \in \{Benign, G3, G4\}$ for all ROIs within a core, where all or at least two of these patterns can happen at the same time; (2) The measured length of cancerous tissues ($\mathcal{L}en$) in a typical core length ($\mathcal{L}en^{typical}$) of 18.0 mm. We propose a new approach for ground-truth probability vector generation, enabling the soft labeling instead of the traditional label encoding methods. For this purpose, using $\mathcal{D}_{train}^{cancer}$ the output of sequence learning step $\mathbf{h}$ is fed into a k-way softmax function, which produces

a probability distribution over the k possible class labels ($k = 3$). Suppose $\mathcal{L}en^{(i)}$ represents the length of cancer for the core that $\mathbf{x}^{(i)}$ belongs to. The ground-truth probability vector of the i^{th} ROI is defined as $\mathbf{p}^{(i)} = [p_1^{(i)}, ..., p_k^{(i)}]$. To estimate these probabilities we define the normalized cancer percentage as $\mathbf{C}^{(i)} = \mathcal{L}en^{(i)}/\mathcal{L}en^{typical}$ ($\mathbf{C}^{(i)} \in [0, 1]$). For $k = 3$:

$$\mathbf{p}^{(i)} = \left[p_1^{(i)} = (1 - \mathbf{C}^{(i)}), p_2^{(i)} = \omega \times \mathbf{C}^{(i)}, p_3^{(i)} = (1 - \omega) \times \mathbf{C}^{(i)}\right] \tag{2}$$

where ω is the cancer regularization factor to control the inherent ratio between pattern G3 and G4 in a way that for the cores with GS $3+4$ label, ω be greater than 0.5 to imply a higher probability of having pattern G3 than the G4 and vice-versa. For ROIs which originate from the cores with GS $3+3$ or $4+4$ readings, ω is set to 1 and 0, respectively. The cost function to be minimized is defined as:

$$J = \frac{1}{|\mathcal{D}_{train}^{cancer}|} \sum_{i=1}^{N} \sum_{k=1}^{K} (\mathbf{p}_k^{(i)} - \overline{\mathbf{p}}_k^{(i)})^2 \tag{3}$$

where $\overline{\mathbf{p}}^{(i)} = [\overline{p}_1^{(i)}, ..., \overline{p}_k^{(i)}]$ is the predictive probability vector.

3.2 Cancer Grading and Tumor in Core Length Estimation

Suppose $\mathcal{C} = \{(\mathbf{x}^{(i)}, y_b^{(i)})\}_{i=1}^{|\mathcal{C}|}$ represent the collection of all labeled ROIs surrounding a target core, where $\mathcal{C} \in \mathcal{D}_{test}$, $|\mathcal{C}| = 80$, $\mathbf{x}^{(i)}$ represents the i^{th} TeUS sequence of the core, and $y_b^{(i)}$ indicates the corresponding binary label. Using the probability output of the first stage model for each ROI, we assign a binary label to each target core. The label is calculated using a majority voting based on the predicted labels of all ROIs surrounding the target. We define the predicted label for each ROI, $\overline{y_b}^{(i)}$, as 1, when $P(\mathbf{y}_b^{(i)}|\mathbf{x}^{(i)}) \geq 0.5$, and as 0 otherwise. The probability of a given core being cancerous based on the cancerous ROIs within that core is $\mathcal{P}_b = \sum_{i=1}^{|\mathcal{C}|} I(\overline{\mathbf{y}}^{(i)} = 1)/|\mathcal{C}|$. A binary label of 1 is assigned to a core, when $\mathcal{P}_b \geq 0.5$. For the cores with prediction of the cancer, we use the output the second stage model to both predict the cancer length and determine a GS for the test core. Suppose $\mathbf{p}_m^{(i)} = \left[p_1^{(i)}, p_2^{(i)}, p_3^{(i)}\right]$ represents the predictive probability output of i^{th} TeUS sequence in the second stage. We define the average predictive probability as $\mathcal{P}^m = \sum_{i=1}^{|\mathcal{C}|} \mathbf{p}_m^{(i)}/|\mathcal{C}|$. Following the histopathology guidelines, to determine a GS for a cancerous test core, $\mathcal{Y}$, we define the core as "GS 4+3 or higher"when $\mathbf{P}_m^{(3)} \geq \mathbf{P}_m^{(2)}$ and otherwise as"GS 3+4 or lower". Furthermore, based on Eq. (2), we can estimate the predicted length of cancer for this core as $\mathcal{L}en^{\mathcal{C}} = (1 - \mathbf{P}_{(1)}^m) \times \mathcal{L}en^{typical}$.

3.3 Model Uncertainty Estimation

We also aim to estimate the model uncertainty in detection of cancer for the areas outside the cancer foci, where the annotation is not available. The key to

estimating model uncertainty is the posterior distribution $P(\Theta|\mathcal{D})$, also referred to a Bayesian inference [9]. Here, we follow the idea in [9] to approximate model uncertainty using Monte Carlo dropout (MC dropout). Given a new input $\mathbf{x}^{(i)}$, we compute the model output with stochastic dropouts at each layer. That is, randomly dropout each hidden unit with certain probability p. This procedure is repeated B times, and we obtain $\{\overline{y}_b^{*(1)}, ..., \overline{y}_b^{*(B)}\}$. Then, the model uncertainty can be approximated by the sample variance, $1/B \sum_{j=1}^{B} (\overline{y}_b^{*(j)} - \hat{y}_b^{*(j)})^2$, where $\hat{y}_b^{*(j)}$ is the average of $\overline{y}_b^{*(j)}$ values.

4 Experiments and Results

Data Division and Model Selection: Data is divided into mutually exclusive patient sets for training, $\mathcal{D}_{train}$, and test, $\mathcal{D}_{test}$. Training data is made up of 80 randomly selected cores from patients with homogeneous tissue regions where the number of cancerous and non-cancerous cores are equal. The test data consists of 170 cores, where 130 cores are labeled as benign, 29 cores with GS $\leq 3+4$, and 12 cores with GS $\geq 4+3$. Given the data augmentation strategy in Sect. 2, we obtain a total number of 80×1,536 = 122,880 training samples ($N = |\mathcal{D}_{train}| = 122,880$). We use 20% of $\mathcal{D}_{train}$ data as the held-out validation sets ($\mathcal{D}_{val}$) to perform the grid search over the number of RNN hidden layers, $n_h \in \{1, 2\}$, batch size, $b_s \in \{64, 128\}$, and initial learning rate, $lr \in \{0.01 - 0.0001\}$, and cancer regularization factor, ω, with three different optimization algorithms, SGD, RMSprop and *Adam*. Results from hyperparameter search demonstrate that network structures with two RNN hidden layers outperform other architectures. Furthermore, for the vanilla RNN, $b_s = 128$, $lr = 0.0001$; for LSTM, $b_s = 64$, $lr = 0.0001$; and for GRU, $b_s = 128$, $lr = 0.01$ generate the optimum models. For all models, $d_r = 0.2$, $l_reg = 0.0001$ generate the lowest loss and the highest accuracy in $\mathcal{D}_{val}$. Also, $\omega = 0.7$ for GS $3+4$ and $\omega = 0.3$ for GS $4+3$ result in the highest performance. After model selection, we use the whole $\mathcal{D}_{train}$ for training a model for the first stage and $\mathcal{D}_{train}^{cancer}$ for the second stage model.

Comparative Method and Baselines: We use standard evaluation metrics as prior approaches [2,4] to quantify our results. We assess the inter-class area under the receiver operating characteristic curve (AUC) for detection of Benign vs. GS $\geq 3+4$ (AUC_1), Benign vs. GS $\geq 4+3$ (AUC_2), and GS $\geq 3+4$ vs. GS $\geq 4+3$ (AUC_3). Table 1 shows the performance comparison between the proposed approach and the following baselines. To substantiate the proposed soft ground-truth label in the second stage of our approach, we replace $\mathbf{p}^{(i)}$ with the labels from multi-label binarization as shown in Fig. 1 (**BL-1**). Also, to justify the necessity of a two-stage approach to tackle the noise, we use the labels from multi-label binarization (**BL-2**) and the weighted version (**BL-3**) in a single stage approach; after the sequence learning step we feed the output to a fully-connected layer with a 3-way softmax function. To generate the weighted version of multi-label binarization labels, for GS $3+4$, the encoded vector is defined as [0, 0.7, 0.3],

Table 1. Model performance for classification of cores in the test data (N = 170)

Method	AUC_1	AUC_2	AUC_3	Average AUC
LSTM	0.96	0.96	**0.86**	**0.93**
GRU	0.92	0.92	**0.84**	**0.89**
Vanilla RNN	0.76	0.76	0.70	0.74
BL-1	0.96	0.96	0.68	0.86
BL-2	0.75	0.68	0.58	0.67
BL-3	0.82	0.84	0.65	0.77
LSTM + GMM-Clustering	0.60	0.74	0.69	0.68
DBN + GMM-Clustering	0.68	0.62	0.60	0.63

and for GS $4+3$ the encoded vector is [0, 0.3, 0.7]. We have also implemented the GMM-clustering method proposed by [4]. We have used the learned feature vector from Deep Belied Network (DBN) method [4] and our best RNN structure (LSTM) to feed the proposed GMM-clustering method. The results suggest that the proposed strategy using both LSTM and GRU cells can lead to a statistically significant improvement in the performance ($p < 0.05$), which is mainly due to a superior performance of our proposed approach in the separation of GS $\geq 3+4$ from GS $\geq 4+3$. It is worthwhile mentioning that core-based approaches like multi-instance learning and traditional multi-class classification are not feasible due to the small number of samples. Also, in the lack of a more clean and reliable dataset, direct modeling of the noise level is not pragmatic [8].

Tumor in Core Length Estimation: Fig. 3 shows the scatter plot of the reported tumor in core length in histopathology vs. the predicted tumor in core length using LSTM cells. The graph shows the correlation between the prediction and histopathology report (correlation coefficient = 0.95). We also calculate the mean squared error (MSE) as the measure of our performance in cancer length estimation where we achieve MSE of 0.12 in the estimation of tumor length.]

Cancer Likelihood Colormaps: Fig. 4(a) shows an example of a cancer likelihood map for biopsy guidance derived from the output of the proposed two-stages approach. Figure 4(b) shows the corresponding estimated uncertainty map generated from the proposed uncertainty estimation method ($p = 0.5$, $B = 100$). Uncertainty is measured as the sample variance for each ROI and normalized to the whole prostate region uncertainty. The level of uncertainty is color-coded using a blue-red spectrum where the blue shows a low level of uncertainty and the dark red indicates the highest level of uncertainty. The uncertainty colormap along with the cancer likelihood map can be used as an effective strategy to harness the possible misguidance during the biopsy.

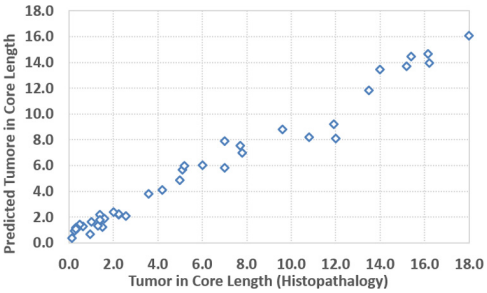
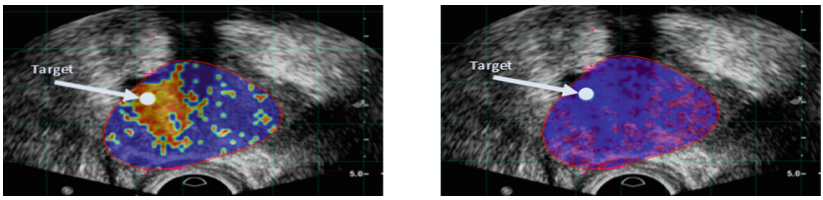


Fig. 3. Scatter plot of the reported tumor in core length in histopathology vs. the predicted tumor in core length



(a) Cancer likelihood map (b) Corresponding uncertainty map

Fig. 4. (a) Cancer likelihood maps overlaid on B-mode image, along with the projected needle path in TeUS data (GS $\geq 4+3$) and centered on the target. ROIs of size 0.5×0.5 mm^2 for which we detect the Gleason grade of 4 and 3 are colored in red and yellow, respectively. The non-cancerous ROIs are colored as blue. (b) The red boundary shows the segmented prostate in MRI projected in TRUS coordinates.[blue=low uncertainty, red=high uncertainty

5 Conclusion

In this paper, we addressed the problem of sparse and noisy histopathology-based ground-truth labels by employing the ground-truth probability vectors as soft labels. These soft labels were estimated by embedding the prior histopathology knowledge about the length of cancer in our two-stage model. The results suggest that soft labels can help the learning process by suppressing the influence of noisy labels and can be used to accurately estimate the length of the suspicious cancer foci. Furthermore, possible misguidance in biopsy is highlighted by the proposed uncertainty measure. Future work will be focused on the analysis of the source of the uncertainty and integrate the proper solution in the framework.

References

1. Ahmed, H.U., et al.: Diagnostic accuracy of multi-parametric MRI and TRUS biopsy in prostate cancer (PROMIS). Lancet **389**(10071), 815–822 (2017)
2. Azizi, S., Bayat, S., Abolmaesumi, P., Mousavi, P., et al.: Detection and grading of prostate cancer using temporal enhanced ultrasound: combining deep neural networks and tissue mimicking simulations. IJCARS **12**(8), 1293–1305 (2017)

3. Azizi, S., et al.: Classifying cancer grades using temporal ultrasound for transrectal prostate biopsy. In: Ourselin, S., Joskowicz, L., Sabuncu, M.R., Unal, G., Wells, W. (eds.) MICCAI 2016. LNCS, vol. 9900, pp. 653–661. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-46720-7_76
4. Azizi, S., Mousavi, P., et al.: Detection of prostate cancer using temporal sequences of ultrasound data: a large clinical feasibility study. Int. J. CARS **11**, 947 (2016). https://doi.org/10.1007/s11548-016-1395-2
5. Azizi, S., et al.: Ultrasound-based detection of prostate cancer using automatic feature selection with deep belief networks. In: Navab, N., Hornegger, J., Wells, W.M., Frangi, A.F. (eds.) MICCAI 2015. LNCS, vol. 9350, pp. 70–77. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-24571-3_9
6. Bayat, S., Azizi, S., Daoud, M., et al.: Investigation of physical phenomena underlying temporal enhanced ultrasound as a new diagnostic imaging technique: theory and simulations. IEEE Trans. UFFC **65**(3), 400–410 (2017)
7. Feleppa, E., Porter, C., Ketterling, J., Dasgupta, S., Ramachandran, S., Sparks, D.: Recent advances in ultrasonic tissue-type imaging of the prostate. In: André, M.P. (ed.) Acoustical imaging, vol. 28, pp. 331–339. Springer, Dordrecht (2007). https://doi.org/10.1007/1-4020-5721-0_35
8. Frénay, B., Verleysen, M.: Classification in the presence of label noise. IEEE Trans. Neural Netw. Learn. Syst. **25**(5), 845–869 (2014)
9. Gal, Y., Ghahramani, Z.: Dropout as a Bayesian approximation: representing model uncertainty in deep learning. In: Machine Learning, pp. 1050–1059 (2016)
10. Kasivisvanathan, V.: Prostate evaluation for clinically important disease: Sampling using image-guidance or not? (PRECISION). Eur. Urol. Suppl. **17**(2), e1716–e1717 (2018)
11. Llobet, R., Pérez-Cortés, J.C., Toselli, A.H.: Computer-aided detection of prostate cancer. Int. J. Med. Inf. **76**(7), 547–556 (2007)
12. Moradi, M., Abolmaesumi, P., Siemens, D.R., Sauerbrei, E.E., Boag, A.H., Mousavi, P.: Augmenting detection of prostate cancer in transrectal ultrasound images using SVM and RF time series. IEEE TBME **56**(9), 2214–2224 (2009)
13. Nelson, E.D., Slotoroff, C.B., Gomella, L.G., Halpern, E.J.: Targeted biopsy of the prostate: the impact of color doppler imaging and elastography on prostate cancer detection and Gleason score. Urology **70**(6), 1136–1140 (2007)
14. Siddiqui, M.M., et al.: Comparison of MR/US fusion-guided biopsy with US-guided biopsy for the diagnosis of prostate cancer. JAMA **313**(4), 390–397 (2015)
15. Singer, E.A., Kaushal, A., et al.: Active surveillance for prostate cancer: past, present and future. Curr. Opin. Oncol. **24**(3), 243–250 (2012)