A Machine Learning Framework for Temporal Enhanced Ultrasound Guided Prostate Cancer Diagnostics

by

Shekoofeh Azizi

B.Sc., Isfahan University of Technology, 2011
M.Sc., Isfahan University of Technology, 2013

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY
in
The Faculty of Graduate and Postdoctoral Studies
(Electrical and Computer Engineering)

THE UNIVERSITY OF BRITISH COLUMBIA
(Vancouver)
March 2020
© Shekoofeh Azizi 2018
The following individuals certify that they have read, and recommend to the Faculty of Graduate and Postdoctoral Studies for acceptance, the thesis entitled:

A Machine Learning Framework for Temporal Enhanced Ultrasound Guided Prostate Cancer Diagnostics

submitted by Shekoofeh Azizi in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Electrical and Computer Engineering.

Examining Committee:

Purang Abolmaesumi, Electrical and Computer Engineering  
Supervisor

Shahriar Mirabbasi, Electrical and Computer Engineering  
University Examiner

Roger Tam, Electrical and Computer Engineering  
University Examiner

Zahra Moussavi, Electrical and Computer Engineering  
External Examiner

Additional Supervisory Committee Members:

Robert Rohling, Electrical and Computer Engineering  
Supervisory Committee Member

Septimiu E. Salcudean, Electrical and Computer Engineering  
Supervisory Committee Member
Abstract

The ultimate diagnosis of prostate cancer involves histopathology analysis of tissue samples obtained through prostate biopsy, guided by either transrectal ultrasound (TRUS), or fusion of TRUS with multi-parametric magnetic resonance imaging. Appropriate clinical management of prostate cancer requires accurate detection and assessment of the grade of the disease and its extent. Despite recent advancements in prostate cancer diagnosis, accurate characterization of aggressive lesions from indolent ones is an open problem and requires refinement.

Temporal Enhanced Ultrasound (TeUS) has been proposed as a new paradigm for tissue characterization. TeUS involves analysis of a sequence of ultrasound radio frequency (RF) or Brightness (B)-mode data using a machine learning approach. The overarching objective of this dissertation is to improve the accuracy of detecting prostate cancer, specifically the aggressive forms of the disease and to develop a TeUS-augmented prostate biopsy system. Towards full-filling this objective, this dissertation makes the following contributions: 1) Several machine learning techniques are developed and evaluated to automatically analyze the spectral and temporal aspect of backscattered ultrasound signals from the prostate tissue, and to detect the presence of cancer; 2) a patient-specific biopsy targeting approach is proposed that displays near real-time cancer likelihood maps on B-mode ultrasound images augmenting their information; and 3) the latent representations of TeUS, as learned by the proposed machine learning models, are investigated to derive insights about tissue dependent features residing in TeUS and their physical interpretation.

A data set consisting of biopsy targets in mp-MRI-TRUS fusion-biopsies with 255 biopsy cores from 157 subjects was used to generate and evaluate the proposed techniques. Clinical histopathology of the biopsy cores was used as the gold-standard. Results demonstrated that TeUS is effective in differentiating aggressive prostate from clinically less-significant disease and non-cancerous tissue. Evidence derived from simulation and latent-feature visualization showed that micro-vibrations of tissue microstructure, captured by low-frequency spectral features of TeUS, is a main source of tissue-specific information that can be used for detection of prostate cancer.
Lay Summary

Prostate cancer is the most frequently diagnosed cancer and the second leading cancer related cause of death in North American men. If detected accurately and managed appropriately, the long-term survival rate is high. The current clinical approach for diagnosis of prostate cancer is through biopsy sampling of the prostate gland and pathological examination of the samples. The biopsy process is guided by ultrasound images to help the physician with selecting the location of tissue samples. The purpose of this thesis is to improve the detection of prostate cancer, especially its aggressive forms, using a new ultrasound technique called Temporal Enhanced Ultrasound (TeUS). This technique overlays additional information about the presence and distribution of prostate cancer on ultrasound images during biopsy, and can help improve the detection of aggressive disease.
Preface

This thesis is primarily based on six journal publications and four conference papers, resulting from a collaboration between multiple researchers and institutes. These publications have been modified to make the thesis coherent. The author was responsible for development, implementation and evaluation of the methods and the production of the manuscripts. All co-authors have contributed to the editing of the manuscripts and providing feedback and comments. Ethical approvals for clinical human studies conducted for this research have been provided by the ethics review board of the National Cancer Institute, National Institutes of Health (NIH) in Bethesda, Maryland.

The data description in Chapter [2] and the study from Chapter [3] is presented at:


Preface

The contribution of the author was in developing the method, implementing and evaluating the method and writing the manuscript. Dr. Mehdi Moradi provided valuable scientific inputs to improve the proposed method for spectral analysis of TeUS. Dr. Farhad Imani extensively contributed in the process of data preparation, generating the target location maps, and preprocessing of TeUS data. Sahar Ghavidel contributed to the proposed spectral feature visualization. Simon Dahonick developed the preliminary version of the codes for overlaying the prostate cancer likelihood maps on TRUS.

The study from Chapter 4 is presented at:


The contribution of the author was in developing the methods, preparation of data, implementing and evaluating the methods and writing the manuscripts.

The study from Chapter 5 is presented at:


The contribution of the author was in developing the methods, implementing and evaluating the methods and writing the manuscripts. Dr. Ming Li acquired the data for the second independent MRI-TRUS fusion biopsy study presented in Chapter 5. Samira Sojoudi and Nathan Van Woudenberg in collaboration with the author developed the software solution that is partly presented in Chapter 5 and Appendix B.

In all of the studies presented in Chapter 3, 4, and 5 Dr. Bradford Wood and Dr. Peter Pinto performed the biopsy procedures, with technical support from Dr. Amir Tahmasebi, Dr. Sheng Xu, Dr. Pingkun Yan and Dr. Jochen Kruecker. Dr. Baris Turkbey and Dr. Peter Choyke provided the radiological readings from mp-MRI for target identification. Dr. Jin Tae Kwak acquired the data at NIH, matched the data to pathology reports and provide us with anonymized, deidentified data for analysis. Dr. Farhad Imani contributed to the data preparation.

The study from Chapter 6 is partly presented at:

Preface


Dr. Sharareh Bayat performed the numerical and TeUS simulation presented in these papers and extensively contributed to the investigation of underlying physical phenomena of TeUS which is presented in Chapter 6 and Appendix A. The author contributed by performing the data preparation, data analysis, statistical investigations and evaluating the results in collaboration with Dr. Sharareh Bayat. Dr. Francois Vignon and Dr. Mohammad Daoud contributed to the physical investigation of TeUS. Dr. Storey Wilson, Dr. Kenneth A. Iczkowski, and Dr. M. Scott Lucia provided the digital pathology whole-mount slides along the gold-standard labels that we used in the simulations. Dr. Guy Nir and Dr. Septimiu Salcudean provided the segmentation of nuclei in those slides. Dr. Larry Goldenberg provided clinical support from the Vancouver Prostate Centre.

Finally, scientific inputs and insight of Prof. Purang Abolmaesumi and Prof. Parvin Mousavi helped with development, implementation and evaluation of all of the proposed methods in the above publications and this thesis. They significantly contributed to editing and improvement of the manuscripts’ structure through their valuable comments and feedback.
Table of Contents

Abstract ................................................................. ii
Lay Summary ............................................................ iii
Preface ................................................................. iv
Table of Contents ...................................................... viii
List of Tables .......................................................... xiii
List of Figures .......................................................... xv
Glossary ............................................................... xxi
Acknowledgements .................................................... xxiii
Dedication ............................................................... xxiv

1 Introduction ........................................................... 1
  1.1 Prostate Cancer Diagnosis ....................................... 1
  1.2 Background ........................................................ 3
    1.2.1 Ultrasound Techniques for Prostate Cancer Diagnosis .. 3
      1.2.1.1 Texture-based and Spectral Analysis ................. 3
      1.2.1.2 Elastography ........................................... 5
      1.2.1.3 Doppler Imaging ...................................... 6
      1.2.1.4 Ultrasound Time Series Analysis .................... 7
    1.2.2 Machine Learning Approaches ............................. 9
      1.2.2.1 Feature Generation and Classification ............... 9
      1.2.2.2 Hidden Markov Models ................................. 10
      1.2.2.3 Deep Learning Approaches ............................ 11
  1.3 Proposed Solution .............................................. 11
    1.3.1 Objectives ................................................ 13
    1.3.2 Contributions ............................................. 14

viii
# Table of Contents

1.3.3 Thesis Outline ........................................... 17

2 Temporal Enhanced Ultrasound Data ........................................... 21
  2.1 Data Acquisition .............................................. 22
  2.2 Histopathology Labeling .......................................... 23
  2.3 Preprocessing and Region of Interest .................................. 25
    2.3.1 Time-domain Representation of TeUS ......................... 27
    2.3.2 Spectral-domain Representation of TeUS ..................... 27
  2.4 Complementary Second Retrospective TeUS Study ..................... 27

3 Detection of Prostate Cancer Using TeUS ........................................... 29
  3.1 Introduction .................................................. 29
  3.2 Spectral Analysis of TeUS for prostate cancer diagnosis ............... 31
    3.2.1 Background: Deep Belief Networks ............................. 31
      3.2.1.1 Restricted Boltzmann Machine ....................... 31
      3.2.1.2 Deep Belief Network ............................... 33
    3.2.2 Classification Framework Based on DBN ....................... 34
      3.2.2.1 Automatic Feature Learning ....................... 34
      3.2.2.2 Cancer Classification ............................... 35
      3.2.2.3 Spectral Feature Visualization .................... 36
    3.2.3 Results and Discussion ..................................... 37
      3.2.3.1 Data Division ...................................... 37
      3.2.3.2 Hyper-parameter Selection ............................ 38
      3.2.3.3 Classification Performance ............................ 39
      3.2.3.4 Choice of Training Data .............................. 40
      3.2.3.5 Colormaps .......................................... 41
      3.2.3.6 Analysis of Tumor Size .............................. 41
      3.2.3.7 Feature Visualization .............................. 43
  3.3 Temporal Analysis of Temporal Enhanced Ultrasound .................... 44
    3.3.1 Background: Recurrent Neural Networks ........................ 44
    3.3.2 Classification Framework Based on RNN ........................ 47
      3.3.2.1 Proposed Discriminative Method ..................... 47
      3.3.2.2 Cancer Classification ............................... 48
    3.3.3 Experiments .............................................. 49
      3.3.3.1 Data Division ...................................... 49
      3.3.3.2 Hyper-parameter Selection ............................ 49
      3.3.3.3 Model Training and Evaluation ....................... 51
      3.3.3.4 Implementation .................................... 52
    3.3.4 Results and Discussion ..................................... 52
## Table of Contents

3.3.4.1 Model Selection ........................................ 52
3.3.4.2 Model Performance ..................................... 53
3.3.4.3 Comparison with Other Methods ....................... 54
3.3.4.4 Network Analysis ....................................... 55
3.4 Conclusion ................................................ 57

### 4 Detection of High-Grade Prostate Cancer Using TeUS ... 59
4.1 Introduction ................................................ 59
4.2 Prostate Cancer Grading Using Spectral Analysis of TeUS .. 61
  4.2.1 Materials ............................................... 61
  4.2.2 Method .................................................. 62
    4.2.2.1 Feature Learning .................................. 62
    4.2.2.2 Distribution Learning ............................... 64
    4.2.2.3 Prediction of Gleason Score ....................... 64
  4.2.3 Results and Discussion .................................. 65
    4.2.3.1 Prostate Cancer Detection and Grading ............ 66
    4.2.3.2 Integration of TeUS and mp-MRI ................... 67
    4.2.3.3 Sensitivity Analysis ................................ 68
4.3 Temporal Analysis of TeUS for prostate cancer grading ... 71
  4.3.1 Method .................................................. 73
    4.3.1.1 Discriminative Model ............................... 73
    4.3.1.2 Cancer Grading and Tumor in Core Length
      Estimation .............................................. 75
    4.3.1.3 Model Uncertainty Estimation ..................... 76
  4.3.2 Experiments and Results ................................ 76
    4.3.2.1 Data Division and Model Selection ................ 76
    4.3.2.2 Comparative Method and Baselines ................. 77
    4.3.2.3 Tumor in Core Length Estimation .................. 78
    4.3.2.4 Cancer Likelihood Colormaps ..................... 78
4.4 Conclusion ................................................ 79

### 5 Decision Support System for Prostate Biopsy Guidance ... 81
5.1 Introduction ................................................ 81
5.2 Transfer Learning From TeUS RF to B-mode Spectral Features ... 83
  5.2.1 Materials ............................................... 84
    5.2.1.1 Unlabeled Data .................................... 84
    5.2.1.2 Labeled Data ...................................... 84
  5.2.2 Methods ................................................ 84
    5.2.2.1 Unsupervised Domain Adaptation ................... 86
    5.2.2.2 Supervised Classification ......................... 88
Table of Contents

5.2.2.3 Baseline Classification .......................... 88
5.2.2.4 Generalization ............................ 88
5.2.3 Results and Discussion ............................ 88
  5.2.3.1 Unsupervised Domain Adaptation ................. 89
  5.2.3.2 Supervised Classification ....................... 90
  5.2.3.3 Baseline Classification ......................... 92
  5.2.3.4 Generalization .............................. 93
  5.2.3.5 Colormaps: .................................. 95
5.3 Transfer Learning From TeUS RF to B-mode Using RNN .... 95
  5.3.1 Materials ....................................... 96
    5.3.1.1 Data Division ............................. 97
    5.3.1.2 Complementary Second Retrospective Study 97
  5.3.2 Methods ........................................ 98
  5.3.3 Results and Discussion .......................... 99
    5.3.3.1 Classification model validation ............... 99
    5.3.3.2 System assessment ........................ 100
  5.3.4 Discussion and comparison with other methods ....... 101
5.4 Conclusion ....................................... 102

6 Investigation of Physical Phenomena Underlying TeUS .... 104
  6.1 Introduction ..................................... 104
  6.2 Spectral Feature Visualization ....................... 106
    6.2.1 Materials ..................................... 106
    6.2.2 Methodology .................................. 106
  6.3 Histopathology Mimicking Simulation ................. 108
    6.3.1 Digital Pathology Data ....................... 108
    6.3.2 Numerical Simulation Design .................. 108
    6.3.3 TeUS Simulation ............................. 109
  6.4 Experiments and Results ........................... 110
    6.4.1 Feature Visualization Results ................. 110
    6.4.2 Simulation Results ........................... 111
  6.5 Conclusion ....................................... 112

7 Conclusion and Future Work ........................... 113
  7.1 Conclusion and Summary ............................ 113
  7.2 Future Work and Suggestions for Further Development .... 117

Bibliography ......................................... 119
# Table of Contents

## Appendices

A Theoretical Background of Temporal Enhanced Ultrasound [140]

B TeUS Biopsy Guidance System Implementation [143]
  B.1 TeUS biopsy guidance system [143]
    B.1.1 TeUS-client [144]
    B.1.2 TeUS Server [146]
## List of Tables

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>Details of equipment and imaging parameters used for TeUS data collection.</td>
<td>23</td>
</tr>
<tr>
<td>2.2</td>
<td>Gleason score distribution in the first retrospective TeUS dataset.</td>
<td>25</td>
</tr>
<tr>
<td>2.3</td>
<td>Gleason score distribution in the second retrospective TeUS study.</td>
<td>28</td>
</tr>
<tr>
<td>3.1</td>
<td>Gleason score distribution in TeUS test and train dataset. Table shows the number of cores for each category.</td>
<td>38</td>
</tr>
<tr>
<td>3.2</td>
<td>Model performance for classification of testing cores for different MR suspicious levels. N indicates the number of cores in each group.</td>
<td>40</td>
</tr>
<tr>
<td>3.3</td>
<td>Model performance in the fold validation analysis for testing cores in datasets $D^A_{test}$ and $D^B_{test}$.</td>
<td>41</td>
</tr>
<tr>
<td>3.4</td>
<td>Model performance for classification of cores in the test data (N = 171).</td>
<td>54</td>
</tr>
<tr>
<td>3.5</td>
<td>Model performance for classification of cores in the test data for Moderate MR suspicious level. N indicates the number of cores in each group.</td>
<td>54</td>
</tr>
<tr>
<td>3.6</td>
<td>Model performance for classification of cores in the test data for High MR suspicious level. N indicates the number of cores in each group.</td>
<td>54</td>
</tr>
<tr>
<td>4.1</td>
<td>Gleason score distribution in TeUS test and train dataset. Table represents the number of cores for each category.</td>
<td>62</td>
</tr>
<tr>
<td>4.2</td>
<td>Model performance for prostate cancer grading in the test dataset using TeUS only and by integration of TeUS and mp-MRI. L is the largest length of the tumor visible in mp-MRI.</td>
<td>66</td>
</tr>
<tr>
<td>4.3</td>
<td>Model performance for classification of cancerous vs. non-cancerous cores in the test dataset using TeUS only and Integration of TeUS and mp-MRI. L is the greatest length of the tumor visible in mp-MRI.</td>
<td>68</td>
</tr>
</tbody>
</table>
List of Tables

4.4 Model performance for classification of cores in the test data (N = 170). AUC₁, AUC₂ and AUC₃ refer to detection of Benign vs. GS≤3+4, Benign vs. GS≥4+3, and GS≤3+4 vs. GS≥4+3, respectively. ............................. 77

5.1 Model performance measured by AUC for classification in different data divisions. ................................................. 92

5.2 Model performance measured by specificity and sensitivity for classification in different data divisions. .................. 92

5.3 Performance for the combination of mp-MRI and TeUS measured by AUC for classification in different data divisions. . . 93

5.4 Performance for the combination of mp-MRI and TeUS measured by specificity and sensitivity for classification in different data divisions. ........................................... 94

5.5 Comparison of model performance measured by AUC using baselines and the proposed approach in different data divisions. 94

5.6 Performance for the TeUS only and combination of mp-MRI and TeUS measured by AUC, specificity, and sensitivity for classification in the test dataset. .............................. 94

5.7 Gleason score distribution in the second retrospective clinical study. ................................................................. 98

5.8 Model performance for classification of cores in the test data from the first retrospective study for different MR suspicious levels. N indicates the number of cores in each group. .... 100

5.9 Run-time of the steps of the prostate guidance system averaged over N = 21 trials with data from the second retrospective study (given as mean±std). ................................. 101
List of Figures

1.1 A schematic diagram of Temporal Enhanced Ultrasound (TeUS) data generation. A time series of a fixed point in an image frame, shown as a red dot, is analyzed over a sequence of ultrasound frames to determine tissue characteristics. 

1.2 A schematic diagram of TeUS-based workflow for prostate biopsy.

2.1 UroNav MR/US fusion system: The identified mp-MRI lesions were delineated on the T2-weighted MR image as the biopsy targets. The target location is shown by the green point along the projected needle path in the ultrasound image.

2.2 Statistics of histopathology and MR readings in our TeUS dataset: Histopathology reports include the Gleason Score (GS) and the percentage distribution of prostate cancer. The MR scores were grouped into three descriptors of “low”, “moderate” and “high”, and referred to as the MR suspicious level assigned to the area.

2.3 Example of distance maps and their corresponding B-mode and RF frames: (a) RF distance map, (b) RF frame, (c) B-mode distance map, (d) B-mode frame. The dark blue is showing the target location and the color spectrum from blue to yellow is showing farther distance from the target.

2.4 Preprocessing and ROI selection: the target region is divided to 80 ROIs of size 0.5 \( mm \times 0.5 \ mm \) and then a sliding window is used for the data augmentation.

3.1 An illustration of a Restricted Boltzmann Machine (RBM): RBM consists of a layer of binary stochastic visible units \( \mathbf{v} \), connected to a layer of stochastic hidden units \( \mathbf{h} \) by symmetrically weighted connections \( \mathbf{W} \). 

8

12

23

24

26

26

31
3.2 An illustration of the proposed method for prostate cancer
detection. Our DBN has a layer of real-valued visible units of
dimension $F = 50$ and four hidden layers with 100, 50 and 6
hidden units. The red box contains the pre-trained DBN, and
the blue box containing one neuron is added for the fine-tuning
step. The latent features are the output of the last layer of
DBN. .................................................. 35

3.3 An illustration of the proposed feature visualization method. 37

3.4 Cancer probability maps overlaid on B-mode ultrasound im-
age, along with the projected needle path in the temporal
ultrasound data and centered on the target. The ROIs for
which the cancer likelihood is more than 70% are colored in
red, otherwise they are colored in blue. The green boundary
shows the segmented prostate in MRI projected in TRUS
coordinates, dashed line shows needle path and the arrow
pointer shows the target: (a) Correctly identified benign core;
(b) Correctly identified the cancerous core. ...................... 42

3.5 Investigation of the effect of tumor size on accuracy. We
obtained the average AUC of 0.77 for cores with MR-tumor-
size smaller than $1.5 \text{ cm}$, and the average AUC of 0.93 for
cores with MR-tumor-size larger than $2 \text{ cm}$. ................. 42

3.6 Differences of distributions between cancerous and benign
tissue back projected in the input neurons: (a) corresponds to
the first neuron in the third hidden layer; (b) corresponds to
the sixth neuron in the third hidden layer. Results are shown
in the frequency range of temporal ultrasound data analyzed
in this section. It is clear that frequencies between $0 - 2 \text{ Hz}$
provide the most discriminative features for distinguishing
cancerous and benign tissue. ................................. 43

3.7 Overview of the proposed method. We use two layers of
RNNs with LSTM cells to model the temporal information
in a sequence of TeUS data. $\mathbf{x}^{(i)} = (x_1, ..., x_T), T = 100$ is
showing the $i^{th}$ sequence data and $x_t$ is indicating the $t^{th}$ time
step. .................................................. 45

3.8 Comparison between optimizer performance for different RNN
cells: Each curve corresponds to an RNN network structure
with two hidden layers, batch size of 128 with the dropout
rate of 0.2 and regularization term of 0.0001. ................. 51
3.9 Learning curves of different RNN cells using the optimum hyper-parameters in our search space. All of the models use the RMSprop optimizer and converge after 65±7 epochs.

3.10 Cancer likelihood maps overlaid on B-mode ultrasound images, along with the projected needle path in the TeUS data, and centered on the target. Red indicates predicted labels as cancer, and blue indicates predicted benign regions. The boundary of the segmented prostate in MRI is overlaid on TRUS data. The arrow points to the target location. The top row shows the result of LSTM and the bottom row shows the result of spectral analysis [1] for benign targets (a), and cancer targets (b) and (c).

3.11 Sequence Length effect: The most discriminative features for detection of prostate cancer can be learned from a fraction of the full TeUS time series.

4.1 An illustration of the proposed cancer grading approach using spectral analysis of TeUS.

4.2 An illustration of the proposed GMM initialization method.

4.3 Cancer likelihood maps overlaid on B-mode US image, along with the projected needle path in the TeUS data and centered on the target. The ROIs for which we detect as Gleason grade of 4 and 3 are colored in red and yellow, respectively. The non-cancerous ROIs are colored as blue. The red boundary shows the segmented prostate in MRI projected in TRUS coordinates and the arrow pointer shows the target.

4.4 Target location and distribution of biopsies in the test data. Light and dark gray indicate central and peripheral zones, respectively. The pie charts show the number of cores and their histopathology. The size of the chart is proportional to the number of biopsies (in the range from 1 to 25), and the colors dark red, light red and blue refer to cores with GS≥ 4+3, GS≤ 3+4 and benign pathology, respectively. The top, middle, and bottom rows depict histopathology results, TeUS prediction, and integration of TeUS and MRI, respectively.


4.6 Model performance for different sizes of training dataset using spectral analysis of TeUS and distribution learning.
4.7 Model performance versus the number of features that we used to generate the final model: (a) For all of the cores; (b) for cores with MR-tumor-size $\geq 2.0$ cm. Decreasing the number of features improves the model performance. 

4.8 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.

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

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

4.11 (a) Cancer likelihood maps overlaid on B-mode US image, along with the projected needle path in the TeUS data (GS $\geq 4 + 3$) and centered on the target. The ROIs of size 0.5 $\times$ 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 and the arrow pointer shows the target.[blue=low uncertainty, red=high uncertainty]

5.1 An illustration of the proposed approach for domain adaptation between RF and B-mode time series data in TeUS framework.

5.2 Learning curve for DBN training based on the cross-entropy: (a) for first hidden layer size, (b) for different learning rates (LR), (c) for different mini-batch size (BS). In a coarse search for the meta-parameters we achieved the lowest cross entropy loss with $n = 44$, $LR = 0.2$, and $BS = 10$.

5.3 Distribution shift from B-mode to RF for the top three features before (top row) and after (bottom row) the shared deep network. The proposed domain adaptation method can effectively align features and reduce the distribution shift in common learned feature space.
## List of Figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.4</td>
<td>Influence of labeled dataset size in classification accuracy: performance</td>
<td>93</td>
</tr>
<tr>
<td></td>
<td>of the method measured by AUC, accuracy, sensitivity and specificity in the</td>
<td></td>
</tr>
<tr>
<td></td>
<td>k-fold cross-validation setting for (a) TeUS RF data and (b) TeUS B-mode</td>
<td></td>
</tr>
<tr>
<td></td>
<td>data.</td>
<td></td>
</tr>
<tr>
<td>5.5</td>
<td>The comparative performance of the proposed method measured by AUC over the</td>
<td>95</td>
</tr>
<tr>
<td></td>
<td>baselines for $D_{test}$.</td>
<td></td>
</tr>
<tr>
<td>5.6</td>
<td>Cancer probability maps overlaid on B-mode US image, along with the projected</td>
<td>96</td>
</tr>
<tr>
<td></td>
<td>needle path in the temporal US data and centered on the target. The ROIs</td>
<td></td>
</tr>
<tr>
<td></td>
<td>for which the cancer likelihood is more than 50% are colored in red,</td>
<td></td>
</tr>
<tr>
<td></td>
<td>otherwise they are colored as blue. The red boundary shows the segmented</td>
<td></td>
</tr>
<tr>
<td></td>
<td>prostate in MRI projected in TRUS coordinates, dashed line shows needle</td>
<td></td>
</tr>
<tr>
<td></td>
<td>path and the arrow pointer shows the target: (a)-(c) Correctly identified</td>
<td></td>
</tr>
<tr>
<td></td>
<td>cancerous core using RF time series data; (b)-(d) Correctly identified</td>
<td></td>
</tr>
<tr>
<td></td>
<td>cancerous core using B-mode time series data.</td>
<td></td>
</tr>
<tr>
<td>5.7</td>
<td>Guidance interface implemented as part of a 3D Slicer module: cancer</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td>likelihood map is overlaid on B-mode ultrasound images. Red indicates</td>
<td></td>
</tr>
<tr>
<td></td>
<td>predicted labels as cancer, and blue indicates predicted benign regions.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>The boundary of the segmented prostate is shown with white and the green</td>
<td></td>
</tr>
<tr>
<td></td>
<td>circle is centered around the target location which is shown in the green</td>
<td></td>
</tr>
<tr>
<td></td>
<td>dot.</td>
<td></td>
</tr>
<tr>
<td>6.1</td>
<td>An illustration of the proposed feature visualization method.</td>
<td>107</td>
</tr>
<tr>
<td>6.2</td>
<td>Pathology mimicking simulations framework.</td>
<td>108</td>
</tr>
<tr>
<td>6.3</td>
<td>ROI selection and nuclei-based scatterer generation process: (a) Sample of</td>
<td>109</td>
</tr>
<tr>
<td></td>
<td>histopathology slide [70], where the red boundary depicts the cancer area;</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(b) digitized slide overlaid on the histopathology slide, where green and</td>
<td></td>
</tr>
<tr>
<td></td>
<td>red areas represent the benign and cancer regions, respectively. The</td>
<td></td>
</tr>
<tr>
<td></td>
<td>selected ROIs are shown by black squares; (c) extracted nuclei positions</td>
<td></td>
</tr>
<tr>
<td></td>
<td>in the selected ROIs; left: a cancer region, right: a benign region; (d)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>the extracted positions of nuclei from each ROI is embedded in an FEM</td>
<td></td>
</tr>
<tr>
<td></td>
<td>model.</td>
<td></td>
</tr>
<tr>
<td>6.4</td>
<td>(a) A sample whole-mount histopathology slide of the prostate. Different</td>
<td>110</td>
</tr>
<tr>
<td></td>
<td>regions of cancer and benign tissue are shown in the pathology slide. (b)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>The corresponding simulated B-mode ultrasound image.</td>
<td></td>
</tr>
</tbody>
</table>
List of Figures

6.5  (a) Differences of distributions between cancerous tissues with Gleason patterns 3 and 4 as well as benign tissues back projected in the input neurons corresponds to the first neuron in the third hidden layer; (b) Spectral difference of the simulated TeUS in benign and different cancer tissues. ............... [111]

6.6  (a) Distribution of the power spectrum in the frequency spectrum of simulated TeUS data at the excitation frequency, (b) Distribution of the power spectrum in the frequency spectrum of simulated TeUS data at the first harmonic of the excitation frequency. ......................... [112]

B.1 Overview of the biopsy guidance system. The three steps in the guidance workflow are volume acquisition, classification and guidance. A client-server approach allows for simultaneous and real-time execution of computationally expensive algorithms including TeUS data classification, and prostate boundary segmentation. ................................. [144]

B.2 The software system has a three-tiered architecture. Ovals represent processing elements while arrows show the direction of data flow. In the US machine layer, PLUS is responsible for US data acquisition and communicates with the TeUS-client via the OpenIGTLink protocol. The TeUS client layer includes TeUS guidance, an extension module within the 3D Slicer framework. The TeUS-server layer is responsible for the simultaneous and real-time execution of computationally expensive algorithms and communicates with TeUS-client via the OpenIGTLink protocol. ................................. [145]
Glossary

**AUC** Area Under the ROC Curve

**B-mode** Brightness Mode

**CD** Contrastive Divergence

**CNNs** Convolutional Neural Networks

**DBN** Deep Belief Networks

**DNN** Deep Neural Networks

**DFT** Discrete Fourier Transform

**DRE** Digital Rectal Exam

**DCE** Dynamic Contrast Enhanced

**DWI** Diffusion Weighted Imaging

**EM** ElectroMagnetically

**GMM** Gaussian Mixture Model

**GPU** Graphics Processing Unit

**GRU** Gated Recurrent Unit

**GS** Gleason Score

**HMMs** Hidden Markov Models

**FEM** Finite Element Model

**FWHM** Full Width at Half Maximum

**ICA** Independent Component Analysis

**KL** Kullbac-Leibler
LSTM  Long Short-Term Memory
MRE   Magnetic Resonance Elastography
MRI   Magnetic Resonance Imaging
MSE   Mean Squared Error
mp-MRI multi-parametric MRI
NCI   National Cancer Institute
NIH   National Institutes of Health
PCa   Prostate Cancer
PCA   Principal Component Analysis
PSA   Prostate Specific Antigen
PSF   Point Spread Function
RBF   Radial Basis Function
RBM   Restricted Boltzmann Machine
RF    Radio Frequency
RFE   Recursive Feature Elimination
RNNs  Recurrent Neural Networks
ROC   Receiver Operating Characteristic
ROI   Region of Interest
SEER  Surveillance, Epidemiology, and End Results
SDM   Subspace Disagreement Measure
SGD   Stochastic Gradient Descent
SVM   Support Vector Machine
TeUS  Temporal Enhanced Ultrasound
TRUS  Transrectal Ultrasound
US    Ultrasound
Acknowledgements

First and foremost, I would like to express my sincerest gratitude to Prof. Purang Abolmaesumi, my advisor, for the opportunities you gave to me. Thank you for your trust, patience and invaluable support that opened this chapter of my life. Thanks for the generosity, advice and friendship you offered me throughout my experience at UBC.

I am very grateful for being associated with Prof. Parvin Mousavi, whose insightful advices helped me throughout the course of this endeavor. Her patience, dedication, mentorship and sincere advice in every single step of the thesis have been truly crucial. Thanks for being always supportive and enthusiastic.

I would like to thank supervisory committee Professors Salcudean and Rohling for their valuable comments and advices.

I would like to thank the Natural Sciences and Engineering Research Council of Canada (NSERC), the Canadian Institutes of Health Research (CIHR), Philips Research North America, and UBC for funding this work.
Dedication

I dedicate this to my parents who left fingerprints of grace on my life; my family without whom none of my success would be possible, you are my wings to fly.
Chapter 1

Introduction

If I have seen farther it is by standing on the shoulders of Giants.

— Sir Isaac Newton

1.1 Prostate Cancer Diagnosis

Prostate Cancer (PCa) is a significant public health issue, and approximately 14% of men will be diagnosed with this disease at some point during their lifetime. According to the American and Canadian Cancer Societies, prostate cancer accounts for 24% of all new cancer cases and results in 33,600 deaths per year in North America. If diagnosed in the early stages, it can be managed with a long-term disease-free survival rate above 90%; even if it is identified later, interventions can be used to increase the life expectancy of patients. The prostate cancer-related death rate has declined significantly (almost 4% per annum) between 2001 and 2009 due to improved testing and better treatment options. The majority of the cases diagnosed today are the early-stage disease, where several treatment options are available, including surgery, brachytherapy, thermal ablation, external beam therapy, and active surveillance. Early detection and accurate staging of prostate cancer are essential to the selection of optimal treatment options. Hence, reducing the disease-associated morbidity and mortality.

Currently, prostate cancer detection is carried out by a combination of Digital Rectal Exam (DRE), measurement of the Prostate Specific Antigen (PSA) level, and histological assessment of biopsy samples. DRE is the most common and least expensive way to screen for prostate cancer. However, DRE is only effective for detecting late-stage prostate cancer in the peripheral zone of the gland, and any abnormalities located in other prostate zones cannot be felt. The PSA test measures the blood level of

1National Cancer Institute (NCI): Surveillance, Epidemiology, and End Results (SEER) Cancer Statistics Review

1.1. Prostate Cancer Diagnosis

PSA, a protein that is produced by the prostate gland and can be used as a biological marker for tumors. The elevated levels often indicate the presence of prostate cancer; however, it also increases by inflammation of the prostate gland (prostatitis), and when prostate enlarges with age (benign prostatic hyperplasia). Therefore, a reliable diagnosis cannot be performed by these two procedures [98]. The definite diagnosis of prostate cancer is core needle biopsy, under Transrectal Ultrasound (TRUS) guidance. The biopsy procedure entails systematic sampling of the prostate followed by histopathology examination of the sampled tissue. In systematic biopsy, up to 12 cores are taken from predefined anatomical locations. Conventional Ultrasound (US) imaging is not capable of distinguishing cancerous and normal tissue with high specificity and sensitivity. Therefore, the biopsy procedure is blind to intraprostatic pathology and can miss clinically significant disease. The sensitivity of conventional systematic biopsy under TRUS guidance, for detection of prostate cancer, has been reported to be as low as 40% [39, 56, 125, 133]. Significant improvement of TRUS-guided prostate cancer biopsy is required to decrease the rate of over-treatment for low-risk disease while preventing the under-treatment of high-risk cancer [89].

Several methods have been proposed to alleviate this issue by enabling patient-specific targeting to improve the detection rate of prostate cancer. Fusion of Magnetic Resonance Imaging (MRI) and TRUS-guided biopsy [41, 67, 121] as an emerging technology for patient-specific targeting is being gradually adopted and has shown significant potential for improved cancer yield [31, 118]. A meta-analysis of seven multi-parametric MRI (mp-MRI) studies with 526 patients shows specificity of 0.88 and sensitivity of 0.74, with negative predictive values ranging from 0.65 to 0.94 [37]. Although using mp-MRI in fusion biopsy has resulted in the best clinical results to date, recent studies suggest the high sensitivity of mp-MRI in the detection of prostate lesions but low specificity [3], hence, limiting its utility in detecting disease progression over time [144]. This approach has other limitations: (1) mp-MRI is often unfamiliar to the biopsying clinician; (2) the co-alignment of mp-MRI and TRUS is challenging [80, 81, 148]; and (3) mp-MRI is not specific for detecting prostate cancer with intermediate risk. Moreover, limited accessibility and high expense of MRI make an ultrasound-based prostate cancer detection system more preferable. The advantages of an ultrasound-based system are several folds: TRUS is already accepted as the standard prostate biopsy guidance tool; different ultrasound data and image acquisition techniques are simultaneously available on ultrasound machines, and ultrasound imaging is among the most accessible and least harmful medical imaging approaches.
1.2 Background

1.2.1 Ultrasound Techniques for Prostate Cancer Diagnosis

Since the early 1990s, there have been numerous efforts to improve ultrasound-based tissue typing in the TRUS-guided biopsy. When tissue undergoes ultrasound imaging, returning echoes contain useful information for tissue typing. This information can be applied to discriminate among different tissues or to determine different structures of the same tissue due to diseases such as cancer. For prostate tissue typing, methods that not only distinguish prostate cancer but also provide information on its grade, have the potential to improve the management of cancer and its treatment while preventing over-diagnosis. In this section, we review ultrasound-based tissue typing approaches and their application in prostate cancer diagnosis. Major ultrasound-based tissue typing methods for prostate cancer characterization include texture-based and spectral analysis of ultrasound data, elastography, Doppler, and ultrasound time series analysis.

1.2.1.1 Texture-based and Spectral Analysis

The intensity information in Brightness Mode (B-mode) ultrasound images can be used to differentiate among various tissue types [111, 125]. For texture-based tissue characterization, the image is divided into windows called a Region of Interest (ROI). ROI sizes between 0.1 cm × cm and 1.45 cm × cm have been reported in the literature [111]. Texture-based methods analyze the first and second order statistics of the gray levels of the B-mode images which form a set of features for texture characterization. The first-order statistics include the mean, standard deviation, skewness and kurtosis of the gray level in each ROI. Moreover, the speckle signal-to-noise ratio, maximum and minimum, and the Full Width at Half Maximum (FWHM) of ROIs have been found useful for prostate tissue typing [139]. These features are sensitive to imaging parameters of ultrasound scanners and dissimilar acoustical properties for tissues with the same pathology have been observed [111].

Probability distribution models (e.g., Rayleigh, Rice, and Nakagami) fitted to the estimated histogram of B-mode image intensities have also been shown to provide useful clinical information for tissue characterization [30, 151]. Tsuni et al. [150] showed that the B-mode image would be affected by the system settings and user operations. They suggested that the Nakagami parametric image provides a comparatively consistent image result when different diagnosticians use different dynamic ranges and system gains. Their
result indicated a better performance of this distribution model compared to other models for prostate cancer diagnostics. Although statistical texture-based features extracted from ultrasound B-mode image are important for tissue typing, they have not been used alone [125]. The combination of texture-based features with features extracted from other ultrasound techniques [106] and other modalities [7] have been shown to be effective.

Second-order statistical features have been introduced to overcome these limitations. These features are related to the spatial properties of the image and are extracted from the co-occurrence and auto-correlation matrices of B-mode image [17, 128, 139, 140]. Although statistical texture-based features extracted from B-mode and envelope detected signals are necessary for tissue typing, they are usually combined with features extracted from other methods for tissue characterization [111]. Combining texture-based and clinical features such as location and shape of the hypo-echoic region can be a promising way to detect prostate cancer. It has been shown that cancerous tissues can be detected with high specificity (about 90-95%) and high sensitivity (about 92-96%) applying the combination of the features [59]. The main shortcomings of texture-based methods are their high dependency on imaging settings of the ultrasound scanner, signal attenuation, dropout, and shadowing.

Some of the tissue-dependent features can only be extracted from Radio Frequency (RF) echo signals before they go through the nonlinear process of envelope detection for B-mode image generation [44, 111, 126]. During the past decades, spectral analysis of RF echo signals constituting a single ultrasound image has been used to improve prostate cancer diagnosis. One of the earliest attempts at using spectral analysis of ultrasound has been by developing an analog filter to break the bandwidth of backscattered ultrasound into three bands [64]. A few years later, Lizzi and Feleppa as the pioneers of spectral analysis of ultrasound signals developed a theoretical framework for the spectral analysis of ultrasound [97]. In this framework, they proposed a method for calibration of the spectrum to account for system-dependent effects. Lizzi et al. categorized tissue structures theoretically based on the spectral properties of ultrasound backscattered signals [96]. Spectral features extracted from the power spectrum of the signals have been used to distinguish between normal and cancerous tissues in the prostate by Feleppa et al. [42, 43]. There are models which represent the relation between spectral features and tissue microstructure theoretically [96]. However, in practice, local spectral noise, and system-dependent effects are challenges for using these techniques [96, 122]. Recently, histoscanning [100, 101], a commercial software that uses features of a single RF frame, has been applied
to characterize prostate cancer. Studies have reported an average sensitivity of 60% and specificity of 66% for 146 patients using this approach \cite{101}. In Section \ref{1.2.2}, we will further explore this category of techniques in more details from a machine learning perspective.

\subsection*{1.2.1.2 Elastography}
Soft tissues tend to exhibit higher deformation than stiffer areas when compression is applied. Elastography is an ultrasound imaging approach that aims to capture tissue stiffness. Since cancerous regions of the tissue are usually stiffer than benign regions, elastography can be helpful for prostate cancer characterization \cite{55,123}. The quasi-static methods and dynamic methods are two main ultrasound elastography categories.

Quasi-static ultrasound elastography or strain elastography of the prostate is based on the analysis of the deformation generated by a static compression of the tissue using TRUS. Krouskop \textit{et al.} \cite{88} analyzed the prostate tissue samples to evaluate the elastic properties of the tissue specimens. These properties were also used by Konig \textit{et al.} \cite{85} for image-guided biopsy of the prostate in a group of 404 patients. In this study, 84% of positive cancer patients were correctly identified \cite{85}.

Improvement in biopsy guidance \cite{25,79,85,129} and prostate cancer identification \cite{25,34,38,52,138} are reported in the literature. However, some well-designed studies did not confirm such results \cite{129,151}. The main limitations of strain elastography include lack of uniform compression over the entire gland, dependency on the operator, penetration issues in large glands, and artifacts due to slippage of the compression plane that can occur in up to 32% of images \cite{25,34}. A water-filled balloon may be placed between the probe and the rectal wall to improve the homogeneity of the deformation and reduce the artifacts \cite{4}.

In dynamic methods, a time-varying force is applied to the tissue; it can be either a short transient mechanical force or an oscillatory force with a fixed frequency \cite{49,65}. Several dynamic methods have been proposed including vibro-elastography \cite{2,66,99,137}, Acoustic Radiation Force Impulsion (ARFI) \cite{120}, and shear wave elastography \cite{23,24}. Vibro-elastography is operator independent and can be used to estimate the stiffness of the tissue using TRUS \cite{2,66,99,137}. Shear wave elastography is the most recent elastography technique, which is based on the measurement of low-frequency shear wave velocity propagating through the tissue \cite{32,49}. While the specificity of shear wave elastography has been reported to be as high as 91%, the reported sensitivity can be relatively low (63%) \cite{78,119}. Other
limitations of shear wave elastography include slow frame rate, limited size of the ROI, delay in image stabilization for each acquisition plane and signal attenuation in enlarged prostates, making the evaluation of the anterior transitional zone difficult or impossible [32].

Most of the current clinical elastography systems are only capable of producing an image that visualizes a single tissue physical parameter, such as stiffness or viscosity, while cancerous tissues are complex and non-uniform and cannot be characterized using only one parameter [7, 55, 106]. To address this limitation, more recently, multi-parametric elastography ultrasound and its combination with multi-parametric MRI have been considered. Mohareri et al. [106] showed the potential of multi-parametric quantitative vibro-elastography in prostate cancer detection for the first time in a clinical study including 10 patients. Ashab et al. [7] also combined multi-parametric MRI with multi-parametric ultrasound, including B-mode and vibro-elastography images. In a study including 36 whole mount histology slides, they examined the potential improvement in cancer detection. The idea of capturing tissue stiffness has been extended to MRI as well, and has been explored as Magnetic Resonance Elastography (MRE) technique [135, 136] for detection of tissue abnormalities. In MRE, an external mechanical excitation is applied to the tissue of interest to induce tissue vibrations and it has been shown to be of value in MRI tissue characterization.

Despite major advancement in the elastography technology, all these approaches are subject to the same intrinsic limitations: “not all cancers are stiff, and all stiff lesions are not cancerous” [32, 136].

1.2.1.3 Doppler Imaging

Doppler imaging is an alternative ultrasound-based technique for detection of pathologic conditions. Doppler-based cancer detection methods take advantage of the neovascularization phenomenon in cancerous tissues; changes in cellular metabolism associated with the development of cancer leads to an increase in blood supply to cancer lesions and therefore to neovascularization in the malignant area [57]. A fundamental problem in Doppler-based cancer detection is that blood is a much weaker scatterer of the ultrasound than the surrounding tissue. Therefore, a considerable frequency shift is required to separate a Doppler signal from the background signal. This challenge normally limits Doppler studies to larger vessels with high blood velocity [111]. Moreover, neovascularization related to prostate cancer is usually at the microvascular level which restricts the applicability of Doppler analysis in prostate cancer detection [113]. Today, color flow images, power Doppler
1.2. Background

Imaging, and contrast-enhanced Doppler are all used for detecting prostate cancer and assisting biopsy procedure.

Initial reports on the application of Doppler imaging for prostate studies date back to 1990 when the conventional color flow imaging was utilized for tissue characterization [158]. In a study including 39 subjects, Potdevin et al. [132] used mean of speed in colored pixels and speed-weighted pixel density to locally discriminate prostate cancer. A report from Arger et al. [6] showed that these two features were not substantially different between diverse pathologic tissues. In another study, Tang et al. [147] selected 54 patients with distinct cancer lesions on ultrasound images. They calculated the density of color pixels from Doppler images and used a t-test to analyze the relationship between the density ratio and malignancy. The obtained sensitivity of 91% in this study should be interpreted along with the large size of ROIs. In general, according to different studies, color Doppler imaging is useful for identifying prostate cancer, but targeted biopsy based on color Doppler imaging alone can miss many areas of cancer [119]. In a study of 120 patients, a biopsy regimen consisting of both sextant and color Doppler-directed biopsies was more sensitive than sextant biopsies alone, but the improvement in cancer detection was minimal [86]. Nelson et al. [119] suggested that the area of hyper-vascularity must be large enough to stand out on the Doppler display. Thus, color Doppler may be more sensitive for detection of clinically significant, high-grade lesions. While contrast enhancing agents can increase the intensity of Doppler signal from the micro-vessels [25, 159], legal and technical difficulties have limited the application of such agents.

1.2.1.4 Ultrasound Time Series Analysis

Temporal Enhanced Ultrasound (TeUS) data is defined as the time series of RF echo signals obtained from a stationary position in a tissue without intentional motion of the transducer or the tissue [111] (Fig. 1.1). Since 2007, temporal enhanced ultrasound data has been explored using a machine learning framework to analyze subtle relative variations between various tissue types. It has been shown that features extracted from these variations, highly correlate with the underlying tissue structure [18]. A fundamental departure from prior methods is to display a likelihood map of the presence of cancer, based on machine learning, instead of using pre-defined thresholds for cancer detection. Machine learning eliminates the need for accurate thresholding of features to identify cancer. Moradi et al. utilized spectral features extracted from ultrasound RF time series signals to distinguish between different animal
1.2. Background

Figure 1.1: A schematic diagram of Temporal Enhanced Ultrasound (TeUS) data generation. A time series of a fixed point in an image frame, shown as a red dot, is analyzed over a sequence of ultrasound frames to determine tissue characteristics.

tissue types [108]. The results of this study demonstrated that temporal ultrasound data are sensitive and specific for tissue typing. Moreover, in an ex vivo study, spectral time series features were applied to differentiate between healthy and cancer tissues in 35 human prostate specimens [107]. The results from this study showed that the features extracted from temporal enhanced ultrasound data are significantly more accurate and sensitive compared to the best texture-based and spectral features for detecting prostate cancer. In an in vivo experiment with six radical prostatectomy patients [73], and in a retrospective study with 158 patients undergoing fusion biopsy [11, 13, 82], accurate prostate cancer detection across grades and patients has been demonstrated. Evidence derived from the experiments to-date suggests that both tissue-related and ultrasound signal-related factors lead to tissue typing information in temporal enhanced ultrasound data. These include the cellular structure of the tissue [108] and the thermal properties of the tissue at micro-scale [36]. For the work presented in this dissertation, TeUS is the imaging modality used for data collection in prostate cancer patients.
1.2. Background

1.2.2 Machine Learning Approaches

From the computer-aided diagnosis point of view, prostate cancer detection using ultrasound imaging, can be automated as a classification or a clustering task. Over the past decades, machine learning methods are employed to automate the process of tissue typing using handcrafted features, and more recently, to automatically derive features that optimally represent the data based on tissue types. In this section, I focus the discussion on these two major approaches for deploying machine learning techniques in prostate cancer diagnosis.

1.2.2.1 Feature Generation and Classification

Previously, manually engineered feature representations extracted from ultrasound data [72, 73, 76] have been used with shallow discriminant models such as linear regression models [42–44], k-Nearest Neighbours (k-NN) classifier [98], support vector machines (SVM) [45, 107], random forests [51, 153], Bayesian classifiers [112] and multi-layer feed-forward perceptron networks [44, 108], to differentiate tissue types [111].

Feleppa et al. [42, 43, 45] extracted the spectral-features for tissue typing in prostate cancer from a line fitted to the power spectrum of a single frame of RF echo signal obtained by sonicating the tissue of interest. In this approach, after using Fourier transform to map the RF signals into the spectral domain, they used a linear regression model to explain the relationship between frequency and amplitude in the power spectrum. The extracted features from the regression model, such as intercept and slope were used for differentiation between tissue types. Llobet et al. proposed the adoption of a k-NN classifier to detect cancerous regions in transrectal ultrasound B-mode images of the prostate [98] where they utilized texture-based feature extracted from the ultrasound images to train their classifier.

Mohareri et al. [106] proposed a novel set of features that obtained from vibro-elastography to classify cancerous and normal tissue and compute a cancer probability map. These features include brightness, tissue strain, absolute value of tissue elasticity, relaxation time, and visco-elastic properties. They used a random forest classification framework to combine multi-parametric features and perform tissue characterization. In a leave-one-patient-out study including 10 patients, they achieved area under the receiver operating characteristic curve (AUC) of 0.8 ± 0.01. Ashab et al. [7] proposed prostate tissue classification based on the combination of multi-parametric MRI and multi-parametric ultrasound which includes texture-based features from B-
1.2. Background

mode images and vibro-elastography features. They used LASSO regression and Recursive Feature Elimination (RFE) for feature reduction and selection. Then, a weighted SVM was used for classification. In a study including 36 whole mount histology slides, they achieved an AUC of 0.81 for cancerous ROIs extracted from the peripheral zone of the prostates [7].

More recently, Imani et al. [73, 75] combined extracted features from discrete Fourier and Wavelet transforms of TeUS data and fractal dimension to characterize in vivo prostate tissue using SVMs. Imani et al. [75] used a joint Independent Component Analysis (ICA) algorithm to select the most informative and independent combination of features extracted from in vivo TeUS signal. Then, they trained an SVM classifier to differentiate between normal and cancer tissue. Ghavidel et al. [51] used spectral features of TeUS data and performed feature selection using random forests to classify lower grade prostate cancer from higher grades. Moradi et al. [108, 112] used fractal dimension with Bayesian classifiers and shallow feed-forward neural networks for ex vivo characterization of prostate tissue. Feleppa et al. [44] also used artificial neural networks to assess the performance of spectral features from RF echo signals collected during TRUS imaging of the prostate.

1.2.2.2 Hidden Markov Models

Hidden Markov models (HMMs) represent sequential data, especially time series, as stochastic processes through capturing repetitive patterns and motifs in the data. HMMs are based on probabilistic dependency and hence are well suited for describing time-varying sequences. Llobet et al. [98] proposed the use of HMMs for tissue characterization in TRUS images collected during prostate biopsies. They modeled the echo intensity values along the biopsy-needle path inside of the scanned tissue. The modeled sequences, in this case, are position-dependent rather than time-dependent. Most recently, analysis of TeUS in the temporal domain using probabilistic methods, such as HMMs, have shown significant promise [114, 117]. Nahlawi et al. [115, 117] used HMMs to model the temporal information of TeUS data for prostate cancer detection. In a limited clinical study including 14 subjects, they identified cancerous regions with an accuracy of 85%. Their results also indicate the importance of order in the characterization of tissue malignancy from TeUS data [114].
1.3. Proposed Solution

1.2.2.3 Deep Learning Approaches

Deep learning algorithms are probabilistic models (networks) composed of many layers that transform input data (e.g., images) to outputs (e.g., disease present/absent) while learning increasingly higher level features \[61, 95\]. In the computer-aided diagnosis literature, Deep Neural Networks (DNN) \[63, 95\] have been recently established as a powerful machine learning paradigm. The main advantage of DNNs is their ability to discover informative representation of data which contributes to feature engineering originally performed solely by human experts. Our group previously examined Convolutional Neural Networks (CNNs) to combine temporal and spatial information from TeUS data to detect high-risk prostate cancer. Imani et al. \[74\] used a machine learning framework consisting of the weighted combination of two CNNs. One network takes the information in the temporal sequence as input while the other network uses the spatial arrangement of the ultrasound data. Then, they use the fusion of mp-MRI and TeUS for characterization of higher grade prostate cancer. In a leave-one-patient-out evaluation including 14 patients, they achieved the Area Under the ROC Curve (AUC) of 0.86 for cancerous regions with GS$\geq$3+3, and AUC of 0.89 for those with GS$\geq$3+4.

Besides being computationally expensive, large-scale data is required for effective deep learning and to ensure proper generalization. Further, the difficulty in explaining the operation of deep neural networks has led to imperceptive and repetitive, yet successful use of deep learning methods as a black-box in many areas. The focus of this thesis is on probabilistic modeling of spectral and temporal aspect of TeUS data using Deep Belief Networks (DBN) \[62\] and Recurrent Neural Networks (RNNs) \[50\], respectively. In addition, methodologies are devised to interpret the internal representations learned by deep models to derive insights about tissue dependent features residing in TeUS.

1.3 Proposed Solution

Despite many years of research to improve prostate cancer diagnosis, conventional ultrasound-guided biopsy has low sensitivity, leading to significant rates of under-treatment and over-treatment. Augmentation the biopsy process with ultrasound-based tissue typing techniques that analyze the spectrum and texture of TRUS, or measure mechanical properties of tissue in response to external excitation is useful. However, individual methods have not shown success in multi-center clinical studies, as it is proven difficult to establish a cancer-specific tissue property consistent across the population \[102\].
1.3. Proposed Solution

Figure 1.2: A schematic diagram of TeUS-based workflow for prostate biopsy.

The focus of this thesis is to advance prostate cancer diagnosis through the development of a deep learning solution for real-time analysis of temporal enhanced ultrasound image. The proposed solution will be trained to automatically analyze backscattered ultrasound signals from the tissue over a multitude of TRUS frames and extract tissue-dependent features to discriminate between different cancer grades. The proposed solution can be deployed as a decision support model for patient-specific targeting during biopsy; it can display cancer likelihood maps on B-mode ultrasound images, in real-time, to indicate the presence of cancer. Moreover, its fusion with MRI, where available, can be used to improve biopsy targeting (Figure 1.2).

A fundamental assumption in previous models for TeUS is the availability of RF data in imaging equipment; however, RF is not available on all commercial US scanners and is usually only provided for research purposes. To overcome the challenge of accessibility of RF data, in the TeUS machine learning framework, we propose to use a transfer learning method [130] to transfer knowledge between RF and B-mode data within an ultrasound scanner. Transfer learning is a method that can compensate for the diversity between datasets by learning from a source dataset (RF time series data) and applying the knowledge to a target dataset (B-mode time series data). This approach exploits the common information between the two domains. Common elements between two data domains can be features of the data or parameters of the models that are built using the data [130]. Deep learning methods, such as DBN, have an essential characteristic that makes them well suited for transfer learning: they can identify abstract features that perform well generalizing across various domains [21, 53]. The model is generated from RF or B-mode TeUS before the biopsy procedure and we can
use the generated model to extract real-time tissue-dependent information from the temporal enhanced ultrasound signals. The proposed solution leverages the standard TRUS imaging pipeline (Figure 1.2), hence does not require modifications to the established clinical workflow. We expect that this technology provides a versatile cross-institutional solution for accurate detection of prostate cancer and patient-specific biopsy targeting.

1.3.1 Objectives

As explained before, the extensive heterogeneity in morphology and pathology of prostate adenocarcinoma are challenging factors in obtaining an accurate diagnosis using different imaging modalities. Providing tissue-specific information during TRUS-guided biopsies can assist with targeting regions with high probability of being cancerous. The analysis of TeUS reveals a difference in response between benign and malignant tissues as reported by several previous studies. The objectives of this thesis are:

- **Objective 1. Detection of prostate cancer using TeUS data.** The first step towards improvement of TRUS-based targeted biopsy is to develop an accurate ultrasound-based tissue typing method. Deep learning techniques are proposed to automatically extract tissue-dependent features from *in vivo* temporal enhanced ultrasound data. These abstract high-level features can be used to distinguish between different prostate tissue types.

- **Objective 2. Detection of high-grade prostate cancer using TeUS data.** The second step towards the improvement of the TRUS-based targeted biopsy is to develop a tissue typing method that can differentiate among grades of the disease and its extent. The ubiquity of noise is an important issue for building computer-aided diagnosis models for prostate cancer biopsy guidance where histopathology data is sparse and not finely annotated. A solution is proposed to alleviate this challenge as a part of the TeUS-based prostate cancer biopsy guidance method.

- **Objective 3. Development of a real-time TeUS-based decision model, and enabling its clinical deployment.** Following successful diagnosis and grading of prostate cancer using deep models, the third objective is to enable its clinical deployment as a tool for improving detection of clinically significant prostate cancer in fusion biopsies. Towards this goal, the ability to remove TeUS spectrum
1.3. Proposed Solution

analysis as a preprocessing step is studied by analyzing data directly in the time domain using Recurrent Neural Networks (RNNs). This can help reduce computation time and accelerate cancer likelihood map generation. Due to limited access to raw ultrasound data, transfer learning is devised to generate comparable performance using B-mode images. Real-time implementation of TeUS-based decision model and its evaluation through retrospective data is another step in realizing this objective. Finally, understanding the physical interaction of TeUS data with tissue is not only of theoretical benefit, but can help with practical decisions in building a decision model.

1.3.2 Contributions

This thesis develops, deploys and evaluates a deep learning framework that uses temporal enhanced ultrasound data to output accurate prostate cancer likelihood maps, overlaid on real-time TRUS images. The proposed solution encapsulates the variability associated with the access of raw ultrasound signals in commercial scanners, and can provide complementary information about the grade and extent of prostate cancer. In the course of achieving the objectives, the following contributions were made:

1. Probabilistic modeling of spectral information in temporal enhanced ultrasound:

   - An automatic feature selection framework is proposed based on DBN for spectral analysis of temporal enhanced ultrasound signals from prostate tissue.
   - A statistically significant improvement in the accuracy of prostate cancer detection is demonstrated compared to previously published studies using spectral features of TeUS signals [74, 76].
   - To determine the characteristics of non-cancerous and cancerous cores in TeUS data and their correlation with learned features, a feature visualization approach is developed. This approach is used to identify the most discriminating frequency components of the time series as learned by the classifier.

2. Probabilistic modeling of temporal aspects of TeUS:

   - TeUS data is extensively and explicitly analyzed in temporal domain using time-dependent probabilistic deep networks for the first time.
1.3. Proposed Solution

- Results indicated that RNN can identify temporal patterns in data that may not be readily observable in spectral domain, leading to significant improvements in detection of prostate cancer.
- Algorithms are developed for in-depth analysis and visualization of high-level latent features of LSTM-based RNN.
- A transformational finding, achieved through this analysis, is that the most discriminating features for detection of prostate cancer can be learned from a fraction of the full TeUS time series. Specifically, in this data, less than 50 ultrasound frames were required to build models that accurately detect prostate cancer. This information can be used to optimize TeUS data acquisition for clinical translation.

3. Classifying prostate cancer grade using probabilistic modeling of spectral aspects of TeUS data:

- A novel approach for grading and detection of high-grade prostate cancer using spectral analysis of TeUS with DBN is proposed.
- This approach could successfully differentiate among aggressive prostate cancer (GS≥4+3), clinically less significant prostate cancer (GS≤3+4), and non-cancerous prostate tissues.

4. Classifying prostate cancer grade and its extent based on temporal modeling of TeUS data:

- A novel approach is devised for grading and detection of high-grade prostate cancer using temporal analysis of TeUS with RNN. By encapsulating proposed ground-truth probability vectors, this solution can also precisely estimate cancer length in biopsy cores.
- The accuracy of tissue characterization is statistically significantly improved as compared to previously published studies [10, 12, 76].
- A novel strategy is proposed for depiction of patient-specific color maps for biopsy guidance including the estimated model uncertainty.

5. Development of real-time TeUS-based decision model:

- A transfer-learning approach is developed to address limited access to raw ultrasound data on commercial scanners, and scanner differences in multi-center settings.
1.3. Proposed Solution

- Near real-time augmentation of live standard 2D ultrasound images with cancer likelihood maps generated from the models is implemented.
- The viability of using B-mode TeUS for cancer detection is demonstrated using retrospective data. The initial assessment indicates that the solution is capable of providing guidance information for prostate biopsy procedures.

6. Investigation of the physical phenomena underlying TeUS:

- A method for visualization and interpretation of the learned features from TeUS data is presented.
- Evidence derived from feature visualization points to low-frequency components of TeUS as the most informative features for tissue classification. These components potentially represent the effect of pulsation on prostate tissue microstructure in form of micro-vibrations.
- The effect of micro-vibration is simulated using a medium with preset elasticity and scatterer locations extracted from 14 whole-mount digital histopathology slides.
- Results demonstrated that the distribution and micro-vibration of scatterers could lead to tissue typing information in TeUS. This finding is a major breakthrough in understanding and technical formulation of TeUS after a decade.

One million patients in North America undergo prostate biopsy annually; 70% of 10 core biopsies return negative while up to 34% of the positive yield are under-graded. The mp-MRI, as the state-of-the-art imaging technique for prostate cancer detection, has low accuracy in identifying cancer with intermediate-risk or low volume [89], and has a high rate of false positives, leading to a significant number of negative cores in current fusion biopsies [131, 133]. Cancer likelihood maps displayed by the proposed solution can help systematic biopsy by improving the cancer yield and decreasing the number of unnecessary biopsies. They also benefit fusion biopsy as they can compensate for inaccurate co-alignment of mp-MRI, improving the accuracy of mp-MRI especially for intermediate risk cancer. Innovative and transparent deep learning solutions will allow the technology to work across institutions without bias towards the image settings, equipment or patient pool in one setting. In the long term, this technology can benefit patients by enabling periodic examination of those under active surveillance by this technology alone.
All men who undergo prostate biopsy can take advantage of the proposed technology by improving early detection of prostate cancer and decreasing the number of unnecessary biopsies. Further, the proposed solution can benefit patients by enabling periodic, affordable, widely available and minimally-invasive examination of those under active surveillance, where it is critical to detect clinically significant prostate cancer early with a minimal number of biopsies.

1.3.3 Thesis Outline

The rest of this thesis is subdivided into six chapters as outlined below:

**Chapter 2: Temporal Enhanced Ultrasound Data**

This chapter describes the TeUS data acquired and analyzed in this thesis. It explains the clinical process of data acquisition and establishing ground truth prior to data analysis. Time-domain and spectral-domain representations of TeUS signals are also described.

**Chapter 3: Diagnosis of Prostate Cancer Using TeUS**

For accurate diagnosis of prostate cancer, in this chapter, both the spectral and temporal aspect of TeUS data are analyzed. The focus of the chapter is to distinguish two main categories of the tissues (i.e., cancer and benign) which can be modeled as a binary classification task:

- **Spectral analysis of TeUS using DBN:** In this section, a deep learning approach is proposed to automatically analyze the spectral aspect of temporal ultrasound data obtained from 255 cancer foci identified in mp-MRI. Each target is sampled in axial and sagittal planes. A DBN is trained to automatically learn the high-level latent features of temporal ultrasound data. A support vector machine classifier is then applied to differentiate cancer vs. benign tissue, verified by histopathology. Results indicate that the distance between the biopsy target and the prostate boundary, and the agreement between axial and sagittal histopathology of each target impact the classification accuracy. Using temporal ultrasound data in a fusion prostate biopsy study, a high classification accuracy specifically for moderately scored mp-MRI targets is achieved. These targets are clinically common and contribute to the high false positive rates associated with mp-MRI for prostate cancer detection. Results suggest that temporal enhanced ultrasound data combined with mp-MRI has the potential to reduce the number of unnecessary biopsies in fusion biopsy settings.
1.3. Proposed Solution

- Temporal analysis of TeUS using RNN: In this section, a deep RNN is proposed to explicitly model the temporal information in TeUS. By investigating several RNN models, it is demonstrated that Long Short-Term Memory (LSTM) networks achieve the highest accuracy in detecting prostate cancer in TeUS data. Algorithms are presented for in-depth analysis of LSTM networks. Results suggest that temporal modeling of TeUS using RNN can significantly improve cancer detection accuracy over previously presented spectral analysis of TeUS.

Chapter 4: Detection of High-Grade Prostate Cancer Using TeUS

In this chapter, the focus is on accurate detection of higher grade prostate cancer and the problem of detection of different prostate Gleason grades. This task can be modeled as a probabilistic multi-class classification task. 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. These challenges are addressed through:

- Spectral analysis of TeUS for prostate cancer grading: Histopathological grading of prostate cancer reports the statistics of cancer distribution in a biopsy core. In this section, a coarse-to-fine classification approach, similar to histopathology reporting, is proposed that uses statistical analysis and deep learning to determine the distribution of aggressive cancer in ultrasound image regions surrounding a biopsy target. This approach consists of two steps; in the first step, high-level latent features that maximally differentiate benign from cancerous tissue are learned. In the second step, the statistical distribution of prostate cancer grades in the space of latent features is modeled.

- Temporal analysis of TeUS for prostate cancer grading: In this section, a solution is proposed to alleviate the challenge of noisy and sparse labels in building computer-aided diagnosis models for prostate cancer biopsy guidance. Specifically, prior knowledge from the histopathology is embedded as soft labels in a probabilistic model. These soft labels are used as a replacement for the sparse and noisy labels for the training of an RNN-based model. This information is then applied to detect the grade of cancer and also to estimate the length of cancer in the target.

Chapter 5: Decision Support System for Prostate Biopsy Guidance
In the search for an accurate and practical technique for prostate cancer diagnosis with TeUS in the clinical setting, in this chapter the focus is on two technical challenges: first, the accessibility of raw RF data in commercial scanners and second, real-time processing of RF/B-mode TeUS data using temporal analysis of TeUS with RNN.

- **Transfer learning from RF to B-mode TeUS spectral features**: In a preliminary attempt, a method is presented for prostate cancer detection using TeUS data obtained either from RF ultrasound signals or B-mode images. For the first time, it is demonstrated that by applying domain adaptation and transfer learning methods, a tissue classification model trained on TeUS RF data (source domain) can be deployed for classification using TeUS B-mode data alone (target domain) where both data are obtained on the same ultrasound scanner. This is a critical step for clinical translation of tissue classification techniques that primarily rely on accessing RF data since this imaging modality is not readily available on all commercial scanners in clinics.

- **Transfer Learning from TeUS RF to B-mode Using RNN**: In this section, as a part of an unified software framework for real-time analysis of ultrasound data stream using a deep learning solution, an RNN-based transfer learning solution is proposed. It is demonstrated that prostate cancer detection using near real-time analysis of RF and B-mode TeUS data and deep learning is feasible. The proposed software system allows for depiction of live 2D ultrasound images augmented with patient-specific cancer likelihood maps that have been calculated from TeUS. The performance of the proposed system with data obtained in two independent retrospective clinical studies is evaluated.

**Chapter 6: Investigation of Physical Phenomena Underlying TeUS**

In Chapter 4, TeUS is used to address the problem of grading prostate cancer. The method involves capturing high-level latent features of TeUS with a deep learning approach followed by distribution learning to cluster aggressive cancer in a biopsy core. In this hypothesis-generating study, a deep learning based feature visualization method is utilized as a means to obtain insight into the physical phenomenon governing the interaction of temporal ultrasound with tissue. Evidence derived from the deep learning-based feature visualization pointed to low-frequency components of TeUS as the most informative features for tissue classification. These components potentially represent the effect of pulsation on prostate tissue microstructure. As a result, in chapter, mechanical micro-vibrations of scatterers in phantoms
with various scatterer distributions, reflecting benign and cancerous tissue, derived from digital histopathology data are simulated. It is demonstrated that micro-vibrations of scatterers can be captured by low-frequency spectral features of TeUS, similar to \textit{in vivo} results. These observations together with previous results suggest that the distribution and micro-vibration of scatterers can lead to tissue typing information in TeUS.

\textbf{Chapter 7: Conclusion and Future Work}
This chapter includes a short summary of the thesis followed by a discussion on the limitations of the proposed methods and suggestions for future work.
Temporal Enhanced Ultrasound Data

Data will talk to you if you are willing to listen.
— Jim Bergeson

Temporal Enhanced Ultrasound or TeUS is defined as the time series of ultrasound RF or B-mode frames captured from insonification of tissue over time, without intentionally moving the tissue or the ultrasound probe. As it is shown in Fig. 1.1, the tissue response to this prolonged insonification consists of reflected ultrasound echo-intensity values. This chapter presents the TeUS data used for building and evaluation of our tissue typing framework, which is described in Chapter 3 and Chapter 4 for detection and grading of prostate cancer. Then, we deploy the same data to develop the TeUS-based decision system in Chapter 5. In Section 2.1, we describe the data acquisition process followed by a summary of the method used to establish a ground truth for characterizing prostate cancer tissue in Section 2.2. Section 2.3 explains the process of Region of Interest (ROI) selection and data augmentation. Finally, we describe the temporal and spectral representation of TeUS data in Section 2.3.1 and Section 2.3.2 respectively. Also, for further evaluation of our method, we performed a second case study that we explain in Section 2.4.

2.1 Data Acquisition

Data were obtained with a Philips iU22 ultrasound scanner in fusion prostate biopsy procedures where the biopsy target locations were identified using mp-MRI information, and the biopsy was guided by TRUS. The study was approved by the ethics review board of the National Cancer Institute (NCI), National Institutes of Health (NIH) in Bethesda, Maryland. One hundred and thirty-two (132) enrolled subjects provided informed consent to participate.

Every subject underwent preoperative mp-MRI examination with three pulse sequences: T2-weighted, Diffusion Weighted Imaging (DWI), and Dynamic Contrast Enhanced (DCE) imaging. Before the biopsy procedure, suspicious lesions were identified using mp-MRI (Fig. 2.1). Two independent highly experienced genitourinary radiologists (B.T. and P.L.C.) with 8 and 14 years of experience, interpreted and scored suspicious lesions according to a previously published protocol [152]. Each radiologist assigned an overall score in the range from 1 (no cancer) to 5 (aggressive cancer) to a suspicious area. The consensus scores were grouped into three descriptors of “low” (score of \( \leq 2 \)), “moderate” (score of 3) and “high” (score of \( \geq 4 \)), and referred to as the MR suspicious level assigned to the area. These scores are based on findings on each mp-MRI sequence using previously described criteria [143], which indicate both the presence of prostate cancer and tumor grade. The standardized PI-RADS criteria were not in use for this study.

At the beginning of the procedure, a 3D ultrasound volume of the prostate was reconstructed by obtaining a series of Electromagnetically (EM) tracked 2D TRUS images. The identified mp-MRI lesions were delineated on the T2-weighted MR image as the biopsy targets. Then, using UroNav MR-US fusion system (Invivo Inc., a subsidiary of Philips Healthcare), T2-weighted MR images were registered with the 3D TRUS volume of the prostate [103, 161]. Following the registration of TRUS and MR volumes, the target locations for biopsy were transformed into the EM coordinate frame. A clinician then navigated in the prostate volume towards the MR-identified target. TRUS transducer was held steady for about 5 seconds to acquire 100 frames of TRUS B-mode and RF time series data from the target, followed by firing the biopsy gun to acquire a tissue sample. The Endocavity curved probe (Philips C9-5ec) with the frequency of 6.6 MHz was used for data acquisition. Two cores from axial and sagittal imaging planes were obtained per target location, respectively. Hereafter, we refer to these cores as “axial sample” and “sagittal sample”, respectively. TEUS data was only recorded from the primary lesion in the axial imaging plane to minimize disruption to the clinical workflow. Table 2.1 shows the details of equipment and imaging.
2.2 Histopathology Labeling

Histopathology reports include the Gleason Score (GS) and the percentage distribution of prostate cancer in the axial and sagittal samples from each target. The GS is the most common system to describe the level of abnormality in the prostate tissue. Gleason grades range from 1 (resembling normal tissue) to 5 (aggressive cancerous tissue). Gleason score is reported as a sum

Table 2.1: Details of equipment and imaging parameters used for TeUS data collection.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ultrasound Machine</td>
<td>Philips iU22</td>
</tr>
<tr>
<td>Ultrasound Probe</td>
<td>Philips C9-5ec</td>
</tr>
<tr>
<td>Probe Array</td>
<td>Curved Linear</td>
</tr>
<tr>
<td>Central Frequency</td>
<td>6.6 MHz</td>
</tr>
<tr>
<td>MRI Scanner</td>
<td>3.0-Tesla MRI scanner (Philips Achieva)</td>
</tr>
<tr>
<td>MRI coil</td>
<td>Endorectal coil (BPX-30)</td>
</tr>
</tbody>
</table>

Figure 2.1: UroNav MR/US fusion system: The identified mp-MRI lesions were delineated on the T2-weighted MR image as the biopsy targets. The target location is shown by the green point along the projected needle path in the ultrasound image.
2.2. Histopathology Labeling

Figure 2.2: Statistics of histopathology and MR readings in our TeUS dataset: Histopathology reports include the Gleason Score (GS) and the percentage distribution of prostate cancer. The MR scores were grouped into three descriptors of “low”, “moderate” and “high”, and referred to as the MR suspicious level assigned to the area.

of the grades of the two most common patterns in a tissue specimen [39]. In our dataset, 83 biopsy cores are cancerous with GS 3+3 or higher, where 31 of those are labeled as clinically significant cancer with GS $\geq 4+3$. The remaining 172 cores are non-cancerous and include benign or fibromuscular tissue, chronic inflammation, atrophy and Prostatic Intraepithelial Neoplasia (PIN) [12]. Figure 2.2 and Table 2.2 shows the distribution of different Gleason score and MR reading in the dataset. In our dataset, 70% of the cores have moderate MR suspicious level where 67% of the cores were detected as non-cancerous after the biopsy.
2.3 Preprocessing and Region of Interest

Due to scattering phenomena, ultrasound imaging techniques do not provide accurate information about the location of every individual cells. The echoes reflected from such small objects in soft tissues are scattered in all directions rather than solely in the direction back to the transducer \[60\]. Thus, to generate and to annotate TeUS signal, a group of RF or B-mode values (not a single value), corresponding to areas known as Region of Interest (ROI) is considered. For each biopsy target, we analyze an area of $2 \times 10$ mm $\times$ mm around the target location in the lateral and axial directions, respectively. This area is along the projected needle path in the ultrasound image and centered on the target location. To register the captured target location on B-mode image during biopsy procedure to the corresponding RF image, we use the distance map generated using the scan conversion parameter at NIH. Figure 2.3 shows an example of such a distance map, where the dark blue is showing the target location and the color spectrum from blue to yellow is showing farther distance from the target. We divide this region to 80 equally-sized ROIs of size $0.5 \times 0.5$ mm $\times$ mm (see Fig. 2.4). Considering 1540 m/sec as the propagation speed of ultrasound signal in soft tissue and 3 cycles per pulse, the axial resolution of the system is around 0.2 mm. Each ROI includes 27-55 RF samples based on the depth of imaging and the target location. We also augment the data by creating ROIs using a sliding window of size $0.5 \times 0.5$ mm $\times$ mm over the target region, which results in 1,536 ROIs per target (see Fig. 2.4). We use histopathology outcome to assign a binary $y \in \{0, 1\}$ or non-binary label $y$ to each ROI. We will further discuss the details of ground-truth label assignment and the corresponding obstacles in the coming chapters.
2.3. Preprocessing and Region of Interest

Figure 2.3: Example of distance maps and their corresponding B-mode and RF frames: (a) RF distance map, (b) RF frame, (c) B-mode distance map, (d) B-mode frame. The dark blue is showing the target location and the color spectrum from blue to yellow is showing farther distance from the target.

Figure 2.4: Preprocessing and ROI selection: the target region is divided to 80 ROIs of size $0.5 \text{ mm} \times 0.5 \text{ mm}$ and then a sliding window is used for the data augmentation.
2.3.1 Time-domain Representation of TeUS

Let $C^T = \{(x^{(i)}, y^{(i)})\}^{|C^T|}_{i=1}$ represent the collection of all labeled ROIs surrounding a target core, where $|C^T| = 80$, $x^{(i)}$ represents the $i^{th}$ TeUS sequence of the core, and $y^{(i)}$ indicates the corresponding label. For each ROI, we generate a sequence of TeUS data, $x^{(i)} = (x_1^{(i)}, ..., x_T^{(i)})$, $T = 100$ by averaging over all time series values within a given ROI of an ultrasound frame and subtracting the mean value from the given time series. We call $x^{(i)}$ the time-domain representation of TeUS.

2.3.2 Spectral-domain Representation of TeUS

Let $C^S = \{(f^{(i)}, y^{(i)})\}^{|C^S|}_{i=1}$ represent the collection of all labeled ROIs surrounding a target core, where $|C^S| = 80$, $f^{(i)}$ represents the $i^{th}$ TeUS spectral component of the core, and $y^{(i)}$ indicates the corresponding label. To obtain the spectral-domain representation of TeUS, $f^{(i)}$, we compute the spectrum of TeUS data obtained from each biopsy target and in each region of interest. For this purpose, we take the Discrete Fourier Transform (DFT) of all zero-mean ultrasound time series corresponding to the RF/B-mode samples in each ROI, normalized to the frame rate. Then, we calculate the mean absolute values of the Fourier transforms of the RF/B-mode time series in each ROI. Finally, each ROI is represented by 50 positive frequency components, $f^{(i)} = (f_1^{(i)}, ..., f_{50}^{(i)})$, $F = 50$.

2.4 Complementary Second Retrospective TeUS Study

To further assess the developed solutions in this thesis, we performed a second independent MRI-TRUS fusion biopsy study. This study also follows the exact same protocol as the previous study, however, the dataset includes only TeUS B-mode data. The study was approved by the institutional ethics review board and all subjects provided informed consent to participate. Six subjects were enrolled in the study where they underwent preoperative mp-MRI examination prior to biopsy, to identify the suspicious lesions. Like the previous study, the procedure was performed using UroNav MR-US fusion system and the data were acquired using Philips iU22 ultrasound scanner. From each MRI-identified target locations, two biopsies were taken, one in the axial imaging plane and one in the sagittal imaging plane. Only TeUS B-mode data were recorded for each target to minimize disruption to the
Table 2.3: Gleason score distribution in the second retrospective TeUS study.

<table>
<thead>
<tr>
<th>Gleason Score</th>
<th>Benign</th>
<th>Cancerous</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of Cores</td>
<td>7</td>
<td>14</td>
</tr>
<tr>
<td>Gleason Score</td>
<td>Benign</td>
<td>GS≤3+4</td>
</tr>
<tr>
<td>Number of Cores</td>
<td>7</td>
<td>5</td>
</tr>
<tr>
<td>Gleason Score</td>
<td>Benign</td>
<td>GS 3+3</td>
</tr>
<tr>
<td>Number of Cores</td>
<td>7</td>
<td>2</td>
</tr>
</tbody>
</table>

clinical workflow. We use the histopathology labeling of the cores as the ground-truth to assess the accuracy of the guidance system in detecting the cancerous lesions. This study resulted in 21 targeted biopsy cores with the Gleason score distribution as explained in Table 2.3. These two datasets have different distribution of the cancer grades. The second study includes more cases with higher grade and aggressive lesions. In this study 67% of the cores are cancerous and 43% of the cores are labeled as clinically significant cancer with GS ≥ 4+3. In contrast, as Fig. 2.2 shows, for the first study these ratios are 33% and 12%, respectively.

Extraction of the target location, selection of the region of interest, histopathology labeling, and generation of TeUS representation also has been done with the exact same method that we introduced in the previous section. We use this data, in Chapter 5 and in Section 5.3 to evaluate the proposed RNN-based model for prostate cancer detection using TeUS B-mode data.
Chapter 3

Detection of Prostate Cancer Using TeUS

*There are things I can’t force. I must adjust. There are times when the greatest change needed is a change of my viewpoint.*  
— Denis Diderot

### 3.1 Introduction

The early diagnosis of prostate cancer plays an important role in the choice and the success of treatments [107]. As discussed in Chapter 1, over the past decades, several ultrasound-based techniques have been proposed for characterizing cancerous tissue [111]. The clinical uptake of these methods has been slow, mainly due to the large variability of the tissue characterization results. The primary sources of such variability are the heterogeneous patient population and cancer grades [10]. Moreover, these methods do not generally report clinically sufficient specificity and sensitivity to ensure that all cancer cases can be captured by ultrasound analysis. In the search for a

---

3.1. Introduction

more accurate and pragmatic ultrasound-based solution for prostate cancer diagnosis, in this chapter, we focus on the advancement of probabilistic modeling of temporal enhanced ultrasound as a recent promising tissue typing technique. This chapter is subdivided into two parts aiming to model spectral and temporal aspect of TeUS data by exploiting suitable deep learning approaches. Over the course of probabilistic modeling of TeUS data, we centered our attention to interpretation of the internal representations learned by our models to derive insights about tissue dependent features residing in TeUS.

As we have discussed in Section 1.2.2.1 in previous implementations of temporal enhanced ultrasound technology [72, 82, 153], features were heuristically determined from the spectral/wavelet analysis of the ultrasound data. In addition to challenges associated with defining features that are correlated with the underlying tissue properties, it is also difficult to determine the best combination of those features for effective tissue typing as the number of features increases. The lack of a systematic approach for feature selection can lead to a so-called “cherry picking” of the features [72, 82, 153]. Deep learning approaches have gained significant attention for capturing abstract representations of input data [22] as a replacement for feature selection, and have been successfully used in medical image analysis [94, 156]. In the first part of this chapter, we exploit Deep Belief Networks (DBN) [22] to automatically learn a high-level feature representation from spectral components of temporal ultrasound data that can detect prostate cancer. Subsequently, we use the high-level features in a supervised classification step based on Support Vector Machine (SVM) to generate a cancer likelihood map. Further, we investigate the factors that affect the classification accuracy within a targeted biopsy interface. We demonstrate that this approach is an effective method for identifying both benign and cancerous biopsy cores in TRUS-guided biopsy [13]. Our results indicate that TeUS can complement mp-MR imaging and together they can be an effective tool for cancer detection.

To date, most of the efforts in our group have mainly focused on spectral analysis as the key pre-processing step for feature extraction from TeUS data [11]. Recently, analysis of TeUS in the temporal domain using probabilistic methods, such as Hidden Markov Models (HMMs), have shown significant promise [114, 117]. Thus, in the second part, we propose to use Recurrent Neural Networks (RNNs) [26, 50, 84] to explicitly analyze TeUS data in the temporal domain. Specifically, we use Long Short-Term Memory (LSTM) networks [50] and Gated Recurrent Unit (GRU) [27, 29], the classes of RNNs, to effectively learn long-term dependencies in the data.
3.2 Spectral Analysis of TeUS for prostate cancer diagnosis

3.2.1 Background: Deep Belief Networks

Since the original DBN article was published [63], DBN has become one of the most important models of deep learning. It uses the generative model in the pre-training procedure and back-propagation in the fine-tuning stage [62]. This is very useful when the number of training samples is limited [90], such as medical imaging analysis where the annotation is burdensome and limited to the target regions. In such a case, the unsupervised pre-training of DBN can be beneficial. DBN is also a fast learning algorithm that can find the optimal parameters quicker [62]. Restricted Boltzmann Machines (RBMs) are the building blocks of DBN. In this section, we quickly review the construction and training of DBN.

3.2.1.1 Restricted Boltzmann Machine

An RBM consists of a layer of binary stochastic visible units \( v \), connected to a layer of stochastic hidden units \( h \) by symmetrically weighted connections \( W \).

Figure 3.1: An illustration of a Restricted Boltzmann Machine (RBM): RBM consists of a layer of binary stochastic visible units \( v \), connected to a layer of stochastic hidden units \( h \) by symmetrically weighted connections \( W \).

In an in vivo study with 157 patients, we analyze data from 255 suspicious cancer foci obtained during MRI-TRUS fusion biopsy. We achieve AUC, sensitivity, specificity, and accuracy of 0.96, 0.76, 0.98, and 0.93, respectively. Our results indicate that RNN can identify temporal patterns in the data that may not be readily observable in spectral analysis of TeUS [11], leading to significant improvements in detection of prostate cancer. We also present algorithms for in-depth analysis of LSTM networks.
3.2. Spectral Analysis of TeUS for prostate cancer diagnosis

\( W \) (Fig. 3.1). The joint configuration \((v, h)\) of visible and hidden units has an energy given by:

\[
E(v, h; \theta) = - \sum_{i \in \text{visible}} b_i v_i - \sum_{j \in \text{hidden}} b_j h_j - \sum_{i,j} v_i h_i w_{ij} ,
\]

(3.1)

where \( \theta = \{b_i, b_j, w_{ij}\} \) indicates parameters of the model; \( v_i \) and \( h_j \) are the binary states of visible and hidden units \( i \) and \( j \), \( w_{ij} \) are the weights, \( b_i \) and \( b_j \) are the bias terms. Using this energy function, the network assigns a probability to every possible feature vector at the visible units:

\[
P(v, h; \theta) = \frac{1}{Z(\theta)} \exp(-E(v, h; \theta)) ,
\]

(3.2)

\[
Z(\theta) = \sum_{v} \sum_{h} E(v, h; \theta) ,
\]

(3.3)

where \( Z(\theta) \) is the normalizing constant. The network gives a probability to every input vector via the energy function. The probability of the training vector can be raised by adjusting \( \theta \) to lower the energy. Given a training vector \( v \), the binary states \( h \) of the hidden units follow the conditional distribution:

\[
P(h_j = 1|v) = \sigma(b_j + \sum_i v_i w_{ij}) ,
\]

(3.4)

where \( \sigma(x) = 1/(1 + \exp(-x)) \) is the logistic/sigmoid function. Once binary states of the hidden units \( h \) have been chosen, a reconstruction is produced by setting each \( v_i \) to 1 by following the conditional distribution:

\[
P(v_i = 1|h) = \sigma(b_i + \sum_i h_{j} w_{ij}) ,
\]

(3.5)

The hidden units states are then updated once more so that they represent the features of the reconstruction. The learning of \( w_{ij} \in W \) is done by a method called Contrastive Divergence (CD) [62, 63]. The change in a weight is given by:

\[
\Delta w_{ij} = \epsilon (v_i h_{j \text{data}} - v_i h_{j \text{reconstruction}}) ,
\]

(3.6)

where \( \epsilon \) is a learning rate. The term \( v_i h_{j \text{data}} \) is the fraction of times that the visible unit \( i \) and the hidden units \( j \) are on together when the hidden units are being driven by data. \( v_i h_{j \text{reconstruction}} \) is the corresponding fraction for reconstruction. Through the learning process, we can obtain the optimum
value of $W$. Also, with the same rule, the optimum values of $b_i$ and $b_j \in B$ can be learned.

The power of RBM lies in the form of reconstruction oriented learning. During reconstruction, it only uses the information in hidden units, which is learned as features from the input. If the model can recover original input perfectly, it means that the hidden units retain enough information of the input, and the learned weights and biases can be deemed as good measures of the input data.

### 3.2.1.2 Deep Belief Network

A single hidden layer RBM is not the best way to capture the features in the data. After the training of one RBM, the learned features can be used as the input data for a second RBM. This kind of layer-by-layer learning system can be used to construct DBN. A schematic representation is shown in Fig. 3.2. The DBN is trained in two stages:

1. An unsupervised pre-training phase which sets the weights of the network to the approximately right neighborhood.
2. A fine-tuning phase where the weights of the network are moved to the local optimum by back-propagation on labeled data.

The pre-training is performed from the input layer up to the output layer, following a greedy approach. The pre-training process, as described in Section 3.2.1.1 is repeated several times, layer by layer, obtaining a hierarchical model in which each layer captures strong high-order correlations between its input units. After having greedily pre-trained all network layers [22], the parameters of the deep model, $\{W, B\}$ are then refined using the labeled data and back-propagation algorithm. For this purpose, a final logistic regression layer is added to the end of feature learning system. In particular, the fine-tuning stage minimizes the cross-entropy error as the loss function, $L_{DBN}$:

$$L_{DBN} = -\sum_{i \in \mathcal{O}} o_i \log \hat{o}_i,$$

where $\hat{o}_i$ is the value of the $i_{th}$ unit of the DBN output layer, $\mathcal{O}$ and $o_i$ is the ground-truth value of the corresponding labeled input data, following the one-hot encoding. Each output node is associated to a specific label and in the case of binary classification, we have one final node. Based on the previous formulation, we can see, each layer of RBM in the DBN is a process
3.2. Spectral Analysis of TeUS for prostate cancer diagnosis

of nonlinear feature transformation. Features learned in the top layer of the DBN (before the logistic regression layer) are the most representative features for modeling the data. It can be denoted by \( H_p = (h_{p1}, h_{p2}, ..., h_{pn}) \), where \( p \) represents the top layer, and \( n \) is the number of features in the top layer.

3.2.2 Classification Framework Based on DBN

The objective is to develop a deep learning model to discriminate cancer and benign prostate regions using spectral representation of TeUS data as explained in Section 2.3.2. Let \( D^S = \{ (f^{(i)}, y^{(i)}) \}_{i=1}^{D^S} \) represent the collection of all labeled ROIs, where \( f^{(i)} \) is the \( i^{th} \) TeUS spectral feature vector and \( y^{(i)} \) indicates the corresponding label. An individual TeUS spectral feature of length \( F \), \( f^{(i)} = (f^{(i)}_1, ..., f^{(i)}_j, ..., f^{(i)}_F) \), is composed of \( F = 50 \) frequency components of \( f^{(i)}_j \) and is labeled as \( y^{(i)} \in \{0, 1\} \), where zero and one indicate benign and cancer outcome, respectively in histopathology. We aim to learn a mapping from \( f^{(i)} \) to \( y^{(i)} \) in a supervised framework by using DBN to explicitly model the spectral information in TeUS.

Figure 3.2 shows a schematic diagram of our method. Our framework consists of two parts. First, we use DBN [22] to automatically learn a high-level latent feature representation of the temporal enhanced ultrasound, \( H \), that can detect prostate cancer. Second, we then use the hidden activations of the DBN as the input learned features to an SVM classifier to generate a cancer likelihood map.

3.2.2.1 Automatic Feature Learning

As we explained in Section 3.2.1.2, DBN, as a generative probabilistic model, can be trained to extract a deep hierarchical representation of the input data. We construct a DBN by stacking many layers of RBM, each of which consists of a visible layer (\( v \)) and a hidden layer of neurons (\( h \)). The two layers are connected by a matrix of weighted connections and there are no connections within a layer. RBM can be stacked by linking the hidden layer of one RBM to the visible layer of the next RBM. The role of an RBM is to model the distribution of its input and capturing patterns in the visible units. Thus, DBNs offer an unsupervised way to learn multi-layer probabilistic representations of data that are progressively deeper with each successive layer. Utilizing this representational power, we can find a latent representation of the original low-level spectral features extracted from the temporal enhanced ultrasound. For this purpose, we feed the visible layer,
Figure 3.2: An illustration of the proposed method for prostate cancer detection. Our DBN has a layer of real-valued visible units of dimension $F = 50$ and four hidden layers with 100, 50 and 6 hidden units. The red box contains the pre-trained DBN, and the blue box containing one neuron is added for the fine-tuning step. The latent features are the output of the last layer of DBN.

3.2.2.2 Cancer Classification

Our classification approach is built by an SVM layer to map the learned features from spectrum of TeUS, $f_{DBN} : f \mapsto H_p$, to a posterior over classes. The SVM classification model learns a distribution over classes $P(y|f, \Theta_{DBN}, \Theta_{SVM})$ given a spectral feature vector of $f = (f_1, \ldots, f_F)$, the parameter of the trained DBN, $\Theta_{DBN}$, and the parameter of SVM classifier, $\Theta_{SVM}$.
3.2. Spectral Analysis of TeUS for prostate cancer diagnosis

Θ_{SVM}. The process of training an SVM as a marginal classifier is equivalent to finding the optimal hyper-plane that maximizes the perpendicular distance between the decision boundary and the closest data points in classes and minimizes the decision error on the training dataset [107]. Considering \{H_p^{(1)}, \ldots, H_p^{(N)}\} as training dataset consists of N learned feature vectors with the corresponding label class of \(y_i \in \{-1, 1\}\) (i.e. unlike previous notations, negative one indicate the benign samples), the SVM training problem is equivalent to finding \(\Theta_{SVM} = \{W, b\}\) such that:

\[
\frac{1}{2} \times W^T W + C \sum_{i=1}^{N} \xi^{(i)}, \tag{3.8}
\]

is minimized subject to

\[
y_i(W^T \phi(H_p^{(i)}) + b) \geq 1 - \xi^{(i)}, \tag{3.9}
\]

the slack parameters, \(\xi^{(i)}\), are added to the equations to allow for misclassification of noisy data, and \(C > 0\) is a penalty parameter for the error term that controls over-fitting. The function \(\phi(x)\) maps the data to a higher dimensional space. We use the Radial Basis Function (RBF) kernel, as justified in [107], which is defined as \(K(x, \hat{x}) = \exp(-\gamma \|x - \hat{x}\|^2) = \phi(x)^T \phi(\hat{x})\).

Let \(C^S = \{(f^{(i)}, y^{(i)})\}_{i=1}^{|C^S|}\) represent the collection of all labeled ROIs surrounding a target core, \(|C^S| = 80\), \(f^{(i)}\) represents the \(i^{th}\) TeUS sequence of the core, and \(y^{(i)}\) indicates the corresponding label. Using the probability output of the classifier for each ROI, we assign a binary label to each target core. The label is calculated using a majority vote based on the predicted labels of all ROIs surrounding the target. For this purpose, we define the predicted label for each ROI, \(\overline{y}^{(i)}\), as 1, when \(P_{CS}(y^{(i)}|f^{(i)}) \geq 0.5\), and as 0 otherwise. The probability of a given core being cancerous based on the cancerous ROIs within that core is:

\[
P_{CS} = \frac{\sum_{i=1}^{|C^S|} I(\overline{y}^{(i)} = 1)}{|C^S|}. \tag{3.10}
\]

A binary label of 1 is assigned to a core, when \(P_{CS} \geq 0.5\).

3.2.2.3 Spectral Feature Visualization

To determine the characteristics of the benign and cancerous cores in the temporal enhanced ultrasound data and their correlation with learned features, we propose a feature visualization approach (Fig. 3.3). This approach is used
3.2. Spectral Analysis of TeUS for prostate cancer diagnosis

Figure 3.3: An illustration of the proposed feature visualization method.

to identify the most discriminative features of the time series (i.e. frequency components as introduced in Section 2.3.2), as learned by the classifier. First, data is propagated through the trained DBN and the activations of the last hidden layer, i.e. the learned latent features are computed. To examine the significance of an individual learned latent feature, the activations of all other hidden units in the third layer are set zero. The activation of the non-zero learned feature is back-propagated to the input layer. The resulting signal, displayed in the input layer as a series of frequency components, highlights those components that contribute to the activation of the non-zero learned feature. By comparing the components activated for benign and cancerous cores, we can identify those frequency ranges that are different between two tissue types. This process is performed for all latent features.

3.2.3 Results and Discussion

3.2.3.1 Data Division

The reported registration accuracy for the Philips UroNav MR/US fusion system is $2.4 \pm 1.2mm$ [161]. However, mis-registration is usually more prominent for targets close to the segmented boundary of the prostate. For biopsy cores taken far away from the boundary, we assume that the target is in the center of the core. However, clinicians normally adjust the needle penetration depth for targets that are close to the boundary, especially in the anterior region, so that the core sample is not taken beyond the prostate. To generate our spectral model, we aim to use homogeneous prostate tissue regions with reliable ground-truth labels. Therefore, we select cores for training if they meet all of the following three selection criteria, similar to
Table 3.1: Gleason score distribution in TeUS test and train dataset. Table shows the number of cores for each category.

<table>
<thead>
<tr>
<th>Gleason Score</th>
<th>Benign</th>
<th>GS 3+3</th>
<th>GS 3+4</th>
<th>GS 4+3</th>
<th>GS 4+4</th>
<th>GS 4+5</th>
</tr>
</thead>
<tbody>
<tr>
<td>$D_{\text{train}}^S$</td>
<td>19</td>
<td>0</td>
<td>5</td>
<td>2</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>$D_{\text{test}}^S$</td>
<td>153</td>
<td>26</td>
<td>21</td>
<td>3</td>
<td>17</td>
<td>3</td>
</tr>
</tbody>
</table>

our previous work [13]: 1) located more than 3 \( \text{mm} \) distance to the prostate boundary in TRUS images; 2) have matching histopathology labels between axial and sagittal biopsies; and 3) have a tumor length larger than 7 \( \text{mm} \) if cancerous. We select 32 cores from 29 patients, that fulfill the above criteria, and use the temporal enhanced ultrasound data from them to generate our model. These 32 training cores are labeled as dataset $D_{\text{train}}^S$, where 13 cores are cancerous and 19 cores are benign. we obtain a total number of $32 \times 1,536 = 49,152$ training samples ($N = |D_{\text{train}}^S| = 49,152$) (see Section 2.3). The remaining 223 biopsy cores distributed as presented in Table 3.1 are used as test dataset, $D_{\text{test}}^S$. The test dataset, $D_{\text{test}}^S$, are divided into three sub-groups based on the distance of the target to prostate boundary and agreement between axial and sagittal histopathology labels [124]:

1. Dataset $D_{\text{test}}^A$, consisting of 156 cores from 150 patients whose target distance to prostate boundary ($d$) is $\geq 3 \text{mm}$.

2. Dataset $D_{\text{test}}^B$, consisting of 117 cores from 91 patients in dataset $D_{\text{test}}^A$ that also have the agreement in histopathology labels of axial and sagittal biopsy cores.

3. Dataset $D_{\text{test}}^C$, consisting of 168 cores from 115 patients whose have the agreement in histopathology labels of axial and sagittal biopsy cores.

4. Dataset $D_{\text{test}}^A$, consisting of 67 cores whose target distance to prostate boundary is $< 3 \text{mm}$.

The distribution of the histopathology labels of the $D_{\text{train}}^S$ and $D_{\text{test}}^S$ is summarized in Table 3.1.

### 3.2.3.2 Hyper-parameter Selection

The primary step for training a deep network is to find a proper network structure including the number of hidden layers and the number of neurons
in each layer. Furthermore, there are various hyper-parameters such as the learning rate, the momentum, the weight-cost and the size of each mini-batch that have effects on the training procedure \[61\]. To set these hyper-parameters, we heuristically tried different network structures so that lowest reconstruction error with the default library \[146\] values for all other hyper-parameters is obtained in the training data. Next, we followed the guidelines given by Hinton \[61\] to adjust other hyper-parameters in a way to obtain lower reconstruction error in the training data. The finalized deep network has a real-valued visible layer with \( F = 50 \) units equal to the number of input spectral features. Three hidden layers of the network include 100, 50 and 6 hidden units, respectively. The learning rate, \( \epsilon \), is fixed at 0.001, mini-batch size is 5 and the number of passes is 100 epochs. Moreover, the momentum and the weight cost have the default values of 0.9, and \( 2 \times 10^{-4} \), respectively. The discriminative fine-tuning of the DBN is performed by back-propagation, which requires a final node for representing desired labels of the observations. The fine-tuning step is performed with the learning rate of 0.01 for 70 epochs and mini-batch size of 10. After completion of the learning procedure, the last hidden layer of the DBN produces the latent feature representation. We use training dataset \( D_{S\text{train}} \) to obtain the learned features from the last hidden layer of the trained DBN. Then we use these features as inputs to a non-linear SVM classifier. We have six learned features corresponding to the activations of the six hidden units in the last hidden layer. The SVM classifier uses an RBF kernel; we determine the parameters of the classifier through a grid-search approach \[76\]. For the parameter selection (\( C \) and \( \gamma \)), we exhaustively searched the parameter space \( C \in \{2^{-5}, 2^{-3}, \ldots, 2^5\} \) and \( \gamma \in \{2^{-10}, 2^{-8}, \ldots, 2^4\} \). Following training, we use the SVM classifier on the test data to derive the tissue type labels for each ROI.

### 3.2.3.3 Classification Performance

To assess the performance of our method, sensitivity, specificity, and accuracy were calculated. We also report the overall performance using the area under the receiver operating characteristic curve (AUC). Table 3.2 shows the classification results for test dataset \( D_{S\text{test}} \). Our results indicate that dataset \( D_{B\text{test}} \) has consistently higher classification results than dataset \( D_{A\text{test}} \) across all MR suspicious levels. A closer look at cores in dataset \( D_{A\text{test}} \) also shows that for those samples that are farther from the prostate boundary (at least 5 mm away) and have moderate MR suspicious level (53 cores), we achieve AUC of 0.89, irrespective of mismatch between axial and sagittal histopathology. In
3.2. Spectral Analysis of TeUS for prostate cancer diagnosis

Table 3.2: Model performance for classification of testing cores for different MR suspicious levels. N indicates the number of cores in each group.

<table>
<thead>
<tr>
<th>MR suspicious levels</th>
<th>Dataset $\mathcal{D}_A^{test}$</th>
<th>Dataset $\mathcal{D}_B^{test}$</th>
<th>Dataset $\mathcal{D}_C^{test}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>All MR suspicious levels</td>
<td>156 0.73</td>
<td>117 0.77</td>
<td>168 0.66</td>
</tr>
<tr>
<td>Moderate MR suspicious level</td>
<td>110 0.75</td>
<td>84 0.80</td>
<td>119 0.67</td>
</tr>
<tr>
<td>High MR suspicious level</td>
<td>23 0.78</td>
<td>15 0.71</td>
<td>16 0.84</td>
</tr>
</tbody>
</table>

comparison, only 26% of those cores are identified as cancerous after biopsy which means our approach can effectively complement mp-MRI to reduce the number of false positives for those targets with moderate MR suspicious level.

We also perform a similar analysis for dataset $\mathcal{D}_A^{test}$, where we obtain AUC of 0.36. There are various factors that may have contributed to this drop in classification accuracy, including higher registration error among mp-MRI, TRUS and histopathology for targets close to the segmented prostate boundary, and the inclusion of ultrasound signal from tissue outside the prostate [80, 81]. Moreover, based on the clinical protocol, for targets that are close to the prostate boundary, the biopsy core is not centered on the target location to minimize the penetration of needle in tissue surrounding the prostate. A more accurate ground-truth data needs to be obtained to further validate our approach on targets that are close to the prostate boundary.

Moreover, we perform analysis on dataset $\mathcal{D}_A^{test}$ without the elements of $\mathcal{D}_B^{test}$. This analysis was done on 39 cores from 34 patients whose target distance to the boundary is more than 3 mm and have the disagreement in histopathology labels of axial and sagittal biopsy cores. Our results showed that by using axial plane histopathology as the ground-truth label, we achieved an AUC of 0.73. On the other hand, by using sagittal plane histopathology as the ground-truth label, we obtained an AUC of 0.60. One of the factors that may have contributed to this performance is the fact that temporal ultrasound data is only obtained from the axial plane and no tissue typing information is available from the sagittal plane.

3.2.3.4 Choice of Training Data

In another experiment, to ensure that the classification model does not over-fit to the training data, we trained our SVM classification model using
3.2. Spectral Analysis of TeUS for prostate cancer diagnosis

dataset $\mathcal{D}_{test}^S$ in a fold validation manner. We obtained AUC of 0.71 for $\mathcal{D}_{test}^A$ and AUC of 0.73 for $\mathcal{D}_{test}^B$ in a leave-one-out cross-validation analysis. We also obtained AUC of 0.71 and 0.70 for $\mathcal{D}_{test}^A$ in three-fold and 13-fold cross-validation analysis, respectively. The averaged AUC of leave-one-out cross-validation analysis follows our previous performance results, which supports the generalization of the classification model.

We performed an additional sensitivity analysis by permuting the training and testing data. To create new training and testing sets, in each permutation, we exchanged a randomly selected cancerous or benign core in the training and testing data. The cores are selected from dataset $\mathcal{D}_{test}^B$. This resulted in 32 different permutations given the distribution of cores in our training data. As Table 3.3 shows, on average, we achieved AUC, accuracy, sensitivity, and specificity of 0.70, 71%, 68%, and 70%, respectively.

Table 3.3: Model performance in the fold validation analysis for testing cores in datasets $\mathcal{D}_{test}^A$ and $\mathcal{D}_{test}^B$.

<table>
<thead>
<tr>
<th>Evaluation</th>
<th>Dataset $\mathcal{D}_{test}^A$</th>
<th>Dataset $\mathcal{D}_{test}^B$</th>
</tr>
</thead>
<tbody>
<tr>
<td>leave-one-out cross-validation</td>
<td>0.71 (±0.02)</td>
<td>0.73 (±0.02)</td>
</tr>
<tr>
<td>three-fold cross-validation</td>
<td>0.71 (±0.01)</td>
<td>0.73 (±0.07)</td>
</tr>
<tr>
<td>13-fold cross-validation</td>
<td>0.70 (±0.05)</td>
<td>0.72 (±0.03)</td>
</tr>
</tbody>
</table>

3.2.3.5 Colormaps

Figure 3.4 shows examples of the cancer likelihood maps from dataset $\mathcal{D}_{test}^B$, derived from the output of SVM, overlaid on B-mode ultrasound image. We use the approach described in our earlier publication [107] for this purpose. In the colormaps, red regions belong to ROIs for which the cancer likelihood is more than or equal to 70%. We found that with this threshold, the visualized maps demonstrated all the major tumors in the dataset without a large number of false positives.

3.2.3.6 Analysis of Tumor Size

To investigate the effect of the size of the tumor on our detection accuracy, we analyzed the AUC against the greatest length of the tumor in MRI (ranging from 0.3 cm to 3.8 cm) for $\mathcal{D}_{test}^B$. We obtained the average AUC of 0.77 for cores with MR-tumor-size smaller than 1.5 cm, and the average AUC of 0.93 for cores with MR-tumor-size larger than 2 cm. The results show our method
3.2. Spectral Analysis of TeUS for prostate cancer diagnosis

Figure 3.4: Cancer probability maps overlaid on B-mode ultrasound image, along with the projected needle path in the temporal ultrasound data and centered on the target. The ROIs for which the cancer likelihood is more than 70% are colored in red, otherwise they are colored in blue. The green boundary shows the segmented prostate in MRI projected in TRUS coordinates, dashed line shows needle path and the arrow pointer shows the target: (a) Correctly identified benign core; (b) Correctly identified the cancerous core.

Figure 3.5: Investigation of the effect of tumor size on accuracy. We obtained the average AUC of 0.77 for cores with MR-tumor-size smaller than 1.5 cm, and the average AUC of 0.93 for cores with MR-tumor-size larger than 2 cm. has higher performance for larger tumors. Figure 3.5 shows the effect of the size of the tumor on our detection accuracy and the accuracy of the method when spectral analysis of TeUS is augmented with mp-MR readings.
3.2. Spectral Analysis of TeUS for prostate cancer diagnosis

Figure 3.6: Differences of distributions between cancerous and benign tissue back projected in the input neurons: (a) corresponds to the first neuron in the third hidden layer; (b) corresponds to the sixth neuron in the third hidden layer. Results are shown in the frequency range of temporal ultrasound data analyzed in this section. It is clear that frequencies between $0 - 2 \text{ Hz}$ provide the most discriminative features for distinguishing cancerous and benign tissue.

3.2.3.7 Feature Visualization

For the feature visualization experiment, we found that feature six, corresponding to the hidden activity of the sixth neuron of the third layer, along with features one and four, are those that maximally differentiate cancerous and benign tissue, especially in the lower frequency range. Figure 3.6 shows the visualization of distribution differences for cancerous and benign tissue related to the first and sixth learned features of the third hidden layer, back propagated to the input layer.

While the physical phenomenon governing temporal ultrasound/tissue interaction is the subject of ongoing investigation in our group, several hypotheses have been explored so far. It has been proposed that the acoustic radiation force of the transmit ultrasound signal increases the temperature and changes the speed of sound in different tissue types [36]. It has also been suggested that a combination of micro-vibration of acoustic scatters in microstructures and the density of cells play a role [110]. Our results showed a consistently high classification accuracy in a large dataset in this section. These results suggested that the phenomenon is consistent for the two independent training and test datasets in clinical settings. Interestingly, the range of frequencies that we have identified as most discriminative between
3.3 Temporal Analysis of Temporal Enhanced Ultrasound

cancerous and benign tissue \((0 - 2 \, \text{Hz in Fig.} [3.6])\) are also consistent with
the ranges we have observed in our previous independent studies [75, 76]. We
will further analyze the physical phenomenon governing temporal ultrasound
in Chapter 6 and present the theoretical background of TeUS in Appendix A.

Since DBN is a computationally expensive method, to optimize our
proposed method for real-time display of cancer likelihood maps, a parallel
implementation on Graphics Processing Unit (GPU) is necessary. For such
a parallel implementation, Fourier transform computation, the need for
information transfer between GPU and CPU, and integration of the results
with the UroNav targeted biopsy interface are the bottlenecks. Currently,
the execution time for generating a cancer probability map overlaid on a
B-mode ultrasound image using an Intel Core i7 CPU with 16 GB RAM is
approximately 6 minutes. Thus, an intrinsically time-dependent deep network,
capable of real-time and on-fly analysis of the TeUS frames, is preferable.
Such a model could mitigate the overhead computation of spectral features.
In the next section, we will focus on the temporal analysis of temporal
enhanced ultrasound.

3.3 Temporal Analysis of Temporal Enhanced
Ultrasound

3.3.1 Background: Recurrent Neural Networks

RNNs are a category of neural networks that are “deep” in temporal dimen-
sion and have been used extensively in time-sequence modeling [58]. Unlike a
conventional neural network, RNNs are able to process sequential data points
through a recurrent hidden state whose activation at each step depends on
that of a previous step. Generally, given sequence data \(x = (x_1, ..., x_T)\), an
RNN updates its recurrent hidden state \(h_t\) by:

\[
h_t = \begin{cases} 
0, & \text{if } t = 0 \\
\varphi(h_{t-1}, x_t), & \text{otherwise}
\end{cases}
\]  

(3.11)

where \(x_t\) and \(h_t\) are data values and the recurrent hidden state at time step \(t\),
respectively, and \(\varphi(,\) represents the nonlinear activation function of a hidden
layer, such as a sigmoid or hyperbolic tangent. Optionally, the RNN may
have an output \(y = (y_1, ..., y_T)\). In the traditional RNN model aka vanilla,
the update rule of the recurrent hidden state in (3.11) is implemented as:

\[
h_t = \varphi(Wx_t + Uh_{t-1}),
\]

(3.12)
Figure 3.7: Overview of the proposed method. We use two layers of RNNs with LSTM cells to model the temporal information in a sequence of TeUS data. \( x^{(i)} = (x_1, ..., x_T) \), \( T = 100 \) is showing the \( i^{th} \) sequence data and \( x_t \) is indicating the \( t^{th} \) time step.

where \( W \) and \( U \) are the coefficient matrices of the input at the present step and the recurrent hidden units activation at the previous step, respectively. We can further expand Equation (3.12) to calculate the hidden vector sequence \( h = (h_1, ..., h_T) \):

\[
h_t = \varphi(W_{ih}x_t + W_{hh}h_{t-1} + b_h),
\]

where \( t = 1 \) to \( T \), \( W_{ih} \) denotes the input-hidden weight vector, \( W_{hh} \) represents the weight matrix of the hidden layer, and \( b_h \) is the hidden layer bias vector.

It has been observed that using the traditional RNN implementation, gradients decrease significantly for deeper temporal models. This makes learning of long-term sequence data a challenging task for RNNs. To address this issue, other types of recurrent hidden units such as LSTM and GRU have been proposed. As shown in Equations (3.12) and (3.13), traditional RNN simply applies a transformation to a weighted linear sum of inputs. In contrast, an LSTM-based recurrent layer creates a memory cell \( c \) at each time step whose activation is computed as:

\[
h_t = o_t \varphi(c_t),
\]

where \( o_t \) is the output gate that determines the portion of the memory cell content in time step \( t \) (\( c_t \)) to be exposed at the next time step [50]. The recursive equation for updating \( o_t \) is:

\[
o_t = \sigma(W_{oi}x_t + W_{oh}h_{t-1} + W_{oc}c_{t-1} + b_o),
\]
where $\sigma(.)$ is the logistic sigmoid function, $W_{oi}$ is the input-output weight matrix, $W_{oh}$ is the hidden layer-output weight matrix, and $W_{oc}$ is the memory-output weight matrix. The memory cell, $c_t$, is updated by adding new content, $\overline{c}_t$, and discarding part of the present memory:

$$c_t = i_t \odot \overline{c}_t + f_t \odot c_{t-1} \tag{3.16}$$

where $\odot$ is an element-wise multiplication and $\overline{c}_t$ is calculated as:

$$\overline{c}_t = \varphi(W_{ci}x_t + W_{ch}h_{t-1} + b_c) \tag{3.17}$$

In this equation, the $W$ terms represent weight matrices; e.g., $W_{ci}$ is the input-memory weight matrix. Input gate $i$, and forget gate $f$ determine the degree that new information is to be added and current information is to be removed, respectively, as follows:

$$i_t = \sigma(W_{ix}x_t + W_{ih}h_{t-1} + W_{ic}c_{t-1} + b_i) \tag{3.18}$$

$$f_t = \sigma(W_{fx}x_t + W_{fh}h_{t-1} + W_{fc}c_{t-1} + b_f) \tag{3.19}$$

All weight matrices, $W$, and biases, $b$, are free parameters that are shared between cells across time. Figure 3.7 shows a graphical model of an LSTM cell. A slightly different version of LSTMs are GRUs [29] which have a fewer number of parameters to avoid over-fitting in the lack of sufficient training samples. GRUs combine the forget and input gates into a single update gate, $u$, and merge the cell memory and hidden state to a reset gate, $r$. The activation of $h_t$ of the GRU at time $t$ is a linear interpolation between the previous activation, $h_{t-1}$, and the updated activation, $\overline{h}_t$:

$$h_t = (1 - u_t)h_{t-1} + \overline{h}_tu_t \tag{3.20}$$

where $u_t$, the update gate at time step $t$, determines how much the unit updates its activation or content. The update gate can be calculated as follows:

$$u_t = \sigma(W_{ui}x_t + W_{uh}h_{t-1}) \tag{3.21}$$

where $W_{ui}$ is the input-update weight matrix and $W_{uh}$ denotes the update-hidden weight matrix. The updated activation, $\overline{h}_t$, is computed similarly to the traditional RNN in Equation (3.12) as follows:

$$\overline{h}_t = \varphi(W_{pi}x_t + W_{ph}(r_t \odot h_{t-1})) \tag{3.22}$$
3.3. Temporal Analysis of Temporal Enhanced Ultrasound

Finally, the reset gate, \( r_t \), is computed as:

\[
    u_t = \sigma(W_{ri} x_t + W_{rh} h_{t-1}) .
\]  

(3.23)

3.3.2 Classification Framework Based on RNN

3.3.2.1 Proposed Discriminative Method

Our overarching objective is to develop a deep learning model to discriminate cancer and benign prostate regions in TeUS data. Let \( D^T = \{(x^{(i)}, y^{(i)})\}_{i=1}^{|D^T|} \) represent a collection of all labeled ROIs, where \( x^{(i)} \) is the \( i \)th TeUS sequence and \( y^{(i)} \) indicates the corresponding label. An individual TeUS sequence of length \( T \), \( x^{(i)} = (x^{(i)}_1, ..., x^{(i)}_T) \), is composed of echo-intensity values \( x^{(i)}_t \) for each time step, \( t \), and is labeled as \( y^i \in \{0, 1\} \), where zero and one indicate benign and cancer outcome, respectively in histopathology (see Section 2.3.1). We aim to learn a mapping from \( x^{(i)} \) to \( y^{(i)} \) in a supervised framework by using RNNs to explicitly model the temporal information in TeUS. Our sequence classification approach is built with connected RNN layers followed by a softmax 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 TeUS data. The model learns a distribution over classes \( P(y|x_1, ..., x_T) \) given a time-series sequence \( x_1, ..., x_T \) rather than a single, time independent input. Figure 3.7 shows an overview of the proposed architecture with LSTM cells.

Given the input sequence \( x = (x_1, ..., x_T) \), RNN computes the hidden vector sequence \( h = (h_1, ..., h_T) \) in the sequence learning step. As discussed in Section 3.3.1, \( h \) is a function of the input sequence \( x \), model parameters, \( \Theta \), and time, \( t \): \( \varphi(x; \Theta, t) \). \( \Theta = \{W, B\} \) denotes the sequence learning model parameters, where \( W \) is the set of weights and \( B \) is the set of biases in Eq. (3.13) for vanilla RNN cells, in Eq. (3.14) for LSTM cells, and in Eq. (3.20) for GRU cells through time steps, \( t = 0 \) to \( t = T \). All weight matrices, \( W \), and biases, \( B \), are free parameters that are shared across time. The final node generates the posterior probability for the given sequence:

\[
    z^{(i)} = W_s^T h + b_s ;
\]  

(3.24)

\[
    \hat{y}_b^{(i)} = \arg\max_j S(z_j^{(i)}) , \quad j \in \{0, 1\} , \quad z^{(i)} = W_s^T h + b_s ,
\]  

(3.25)

where \( W_s \) and \( b_s \) are the weight and bias of the fully-connected layer, \( S \) is the softmax function, which in our binary classification case is equivalent to
3.3. Temporal Analysis of Temporal Enhanced Ultrasound

the logistic function, and \( \overline{y}^{(i)} \) indicates the predicted label. The optimization criterion for the network is to maximize the probability of the training labels or equivalently, to minimize the negative log-likelihood defined as a the loss function. This function is the binary cross-entropy between \( y^{(i)} \) and \( \overline{y}^{(i)} \) over all training samples, \( D_{\text{train}}^T = \{ (x^{(i)}, y^{(i)}) \}_{i=1}^N \subset D^T \):

\[
L(\overline{y}, y) = -\frac{1}{N} \sum_{i=1}^N \left[ y^{(i)} \log \overline{y}^{(i)} + (1 - y^{(i)}) \log(1 - \overline{y}^{(i)}) \right],
\]

(3.26)

where \( N = |D_{\text{train}}^T| \). During training, the loss function is minimized through a proper gradient optimization algorithm like stochastic gradient descent (SGD), root mean square propagation (RMSprop) or adaptive moment estimation (Adam) [83].

3.3.2.2 Cancer Classification

The RNN models learn a probability distribution over classes, \( P(y|x_1, \ldots, x_T) \), given a time-series sequence, \( x_1, \ldots, x_T \). Let \( \mathcal{C} = \{ (x^{(i)}, y^{(i)}) \}_{i=1}^{\mathcal{C}^T} \) represent the collection of all labeled ROIs surrounding a target core, where \( \mathcal{C}^T \in D_{\text{test}}^T \), \( |\mathcal{C}| = 80 \), \( x^{(i)} \) represents the \( i \)th TeUS sequence of the core, and \( y^{(i)} \) indicates the corresponding label. Using the probability output of the classifier for each ROI, we assign a binary label to each target core. The label is calculated using a majority vote based on the predicted labels of all ROIs surrounding the target. For this purpose, we define the predicted label for each ROI, \( \overline{y}^{(i)} \), as 1, when \( P(y^{(i)}|x^{(i)}) \geq 0.5 \), and as 0 otherwise. The probability of a given core being cancerous based on the cancerous ROIs within that core is:

\[
P_{\mathcal{C}^T} = \sum_{\mathcal{C}^T}^{\mathcal{C}^T} I(\overline{y}^{(i)} = 1) \frac{|\mathcal{C}^T|}{|\mathcal{C}|}.
\]

(3.27)

A binary label of 1 is assigned to a core, when \( P_{\mathcal{C}^T} \geq 0.5 \).

3.3.2.3 Network Analysis

To better understand the temporal information in TeUS, we examine the LSTM gates. For this purpose, following training, we use the learned weights and biases to regenerate the network behavior for any given sequence of length \( T \). First, the state of each cell is set to zero. Then, the full learning formula (Eqs. (3.14)-(3.19)) along with the model parameters \( \Theta = \{ W, B \} \) are recursively applied for \( T = 100 \) time steps. A summary of the steps is
3.3. Temporal Analysis of Temporal Enhanced Ultrasound

presented in Algorithms 1 and 2. Finally, the on-and-off behavior of the hidden activation in the last layer of the network is used to analyze the high level learned features.

Algorithm 1 Examination of the LSTM gates

**Input:** Trained model parameters “Θ = {W, B}”, input data “X”, number of time-steps “T”, number of input sequence “N”.  

**Output:** States activation “S”, gates activation “G”

Initialization: Set the state of each cell “inStates” to zero.

1: for $i = 0$ to $N$ do  
2: for $t = 0$ to $T$ do  
3: $x ← X(i,:)$  
4: $\{S(i,t), G(i,t)\} ← \text{STEP}(x, \text{inStates}(i,t), W, B)$  
5: $\text{inState}(i,t) ← S(i,t)$  
6: end for  
7: end for  
8: return $S, G$

3.3.3 Experiments

3.3.3.1 Data Division

Data is divided into mutually exclusive training, $\mathcal{D}_{\text{train}}^T$, and test sets, $\mathcal{D}_{\text{test}}^T$. Training data is made up of 84 cores from patients with homogeneous tissue regions (See [14] for more details.). Therefore, the training data are selected from biopsy cores with at least 4.0 mm of cancer in a typical core length of 18.0 mm, 26 of which are labeled as clinically significant cancer with GS $\geq 4+3$. The other half is randomly selected from all available non-cancerous cores. The test data consists of 171 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 Section 2.3 (Fig. 2.4), we obtain a total number of $84 \times 1,536 = 129,024$ training samples ($N = |\mathcal{D}_{\text{train}}^T| = 129,024$).

3.3.3.2 Hyper-parameter Selection

The performance of deep RNNs, similar to other deep learning approaches, are affected by their hyper-parameters. In practice, hyper-parameter selection can be constrained as a generalization-error minimization problem. Solutions are often based on running trials with different hyper-parameter settings, and choosing the setting that results in the best performing model (Fig. 3.8).
3.3. Temporal Analysis of Temporal Enhanced Ultrasound

Algorithm 2 Recurrent Step Function of the LSTM

**Input:** Trained model parameters “\( \Theta = \{W, B\} \)”, input sequence “\( \mathbf{x} \)”, input states of time step \((t-1)\) “\( S \)”.  

**Output:** States activation of the current time step \((t)\) “\( S_t \)”, gates activation of the current time step \((t)\) “\( G_t \)”.

1: procedure \( \text{STEP}(x, S_{t-1}, W, B) \)  
2: \( W_{oi}, W_{oh}, W_{oc}, W_{ci}, W_{ch}, W_{ix}, W_{ih}, W_{fx}, W_{fh}, W_{fc} \leftarrow W \)  
3: \( b_o, b_c, b_i, b_f \leftarrow B \)  
4: \( h_{t-1}, c_{t-1} \leftarrow S_{t-1} \)  
5: \( i_t \leftarrow \sigma(W_{ix}x + W_{ih}h_{t-1} + W_{ic}c_{t-1} + b_i) \)  
6: \( f_t \leftarrow \sigma(W_{fx}x + W_{fh}h_{t-1} + W_{fc}c_{t-1} + b_f) \)  
7: \( o_t \leftarrow \sigma(W_{oi}x + W_{oh}h_{t-1} + W_{oc}c_{t-1} + b_o) \)  
8: \( c \leftarrow \varphi(W_{ci}x + W_{ch}h_{t-1} + b_c) \)  
9: \( c_t \leftarrow i_t \odot c_t + f_t \odot c_{t-1} \)  
10: \( h_t \leftarrow o_t \varphi(c_t) \)  
11: \( S_t \leftarrow \{h_t, c_t\} \)  
12: \( G_t \leftarrow \{i_t, f_t, o_t, c_t\} \)  
13: return \( S_t, G_t \)

optimize the hyper-parameters through a grid search, which is an exhaustive search through a pre-specified subset of the hyper-parameter space of the learning algorithm.

The grid search starts with randomly partitioning the selected training dataset, \( \mathcal{D}_{train}^T \), into training (80%) denoted by \( \mathcal{D}_{tr} \) and held-out validation sets (20%) denoted by \( \mathcal{D}_{val} \). This partitioning results in \( N_{tr} = |\mathcal{D}_{tr}| = 103,219 \) training samples and \( N_{val} = |\mathcal{D}_{val}| = 25,805 \) held-out validation samples. To guide the grid search algorithm, we track the loss, accuracy, and AUC on both \( \mathcal{D}_{tr} \) and \( \mathcal{D}_{val} \). The loss is defined using Eq. (3.26) as the binary cross-entropy between the predicted label and the true label, while accuracy is the percentage of the correctly predicted labels. To stabilize learning and prevent the model from over-fitting on the training data, we use regularization and dropout, two of the most effective proposed strategies [61]. Regularization adds a penalty term to the loss function (Equation (3.26)) to prevent the coefficients from getting too large. Here, we use L2 regularization in the form of \( \lambda \| \omega \|_2^2 \), where we search for \( \lambda \) as the hyper-parameter. Dropout prevents co-adaptations on training data. In each step of training, a dropout layer removes some units of its previous layer from the network, which means the network architecture changes in every training step. These units are chosen randomly based on the probability parameter of the dropout layer as another.
3.3. Temporal Analysis of Temporal Enhanced Ultrasound

Figure 3.8: Comparison between optimizer performance for different RNN cells: Each curve corresponds to an RNN network structure with two hidden layers, batch size of 128 with the dropout rate of 0.2 and regularization term of 0.0001.

hyper-parameter. We perform a 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, $lr \in \{0.01 - 0.0001\}$ with three different optimization algorithms, SGD, RMSprop and Adam [83]. We also experiment with various levels of dropout rate, $d_r \in \{0.2, 0.4\}$ and L2-regularization term ($\lambda$), $l_{reg} \in \{0.0001, 0.0002\}$. These result in 96 different hyper-parameter settings for the proposed approach. All models are trained with the same number of iterations and training is stopped after 100 epochs. Models benefit from reducing the learning rate by a factor once learning stagnates [61]. For this purpose, we monitor the validation loss and if no improvement is observed over 10 epochs, the learning rate is reduced by $l_{new} = l_r \times \text{factor}$, where factor = 0.9.

3.3.3.3 Model Training and Evaluation

Once the optimum hyper-parameters are identified, the entire training set, $D_{train}$, is used to learn the final model. The loss is used as the performance measure for early stopping to avoid overfitting. Training is stopped if the loss as we defined in Eq. (3.26) increases or if it does not decrease after 10 epochs. An absolute change of less than $\delta = 0.0004$ is considered as no improvement in the loss. We also record the model performance for $D_{val}$ to track its behavior in a random subset of training data.

To assess the performance of our method, we report its sensitivity, specificity, and accuracy in detecting cancerous tissue samples in the test data, $D_{test}$. All cancerous target cores are considered as the positive class (labeled as 1), and non-cancerous cores as the negative class (labeled as 0). Sensitivity or recall is defined as the percentage of cancerous cores that are correctly
identified, while specificity is the proportion of non-cancerous cores that are correctly classified. Accuracy is the ratio of the true results (both true positives and true negatives) over the total number of cores. The overall performance of the models is reported using AUC. The curve depicts a relative trade-off between sensitivity and specificity. The maximum value for AUC is 1, where higher values indicate better classification performance.

3.3.3.4 Implementation

We implement the RNNs in Keras [28] using the Tensorflow [1] back-end. Training is performed on a GeForce GTX 980 Ti GPU with 6 GB of memory, hosted by a machine running Ubuntu 16.04 operating system on a 3.4 GHz Intel Core™ i7 CPU with 16 GB of memory. Training of vanilla RNN, LSTM and GRU network structures with 100 epochs takes 1.1, 8.1 and 3.3 hours, respectively. Early stopping and calculation of additional performance metrics are implemented using Keras callbacks and the Tensorflow back-end to evaluate internal states and statistics of the model during training. The proposed network analysis method in Algorithms 1 and 2, is implemented independently of Keras or Tensorflow in Python 2.7, executed on CPU.

3.3.4 Results and Discussion

3.3.4.1 Model Selection

Results from hyper-parameter 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, $d_r = 0.2$ and $lreg = 0.0001$ generate the lowest loss and the highest accuracy for both $D_{tr}$ and $D_{val}$. The learning curves for different optimization algorithms and initial learning rates on $D_{tr}$ are shown in Fig. 3.8. Each curve corresponds to an RNN network structure with two hidden layers, the batch size of 128 with the dropout rate of 0.4 and regularization term of $\lambda = 0.0001$, where the vertical axis is the loss and the horizontal axis is the number of iterations. It is clear that RMSprop substantially outperforms SGD optimization for all of the RNN cell types while RMSprop and Adam have similar performance for GRU and LSTM cells. RMSprop leads to a better performance on our data.

Learning curves of the different RNN cells using the optimum hyper-parameters are shown in Fig. 3.9. The right-vertical axis represents the loss value while the left-vertical axis shows the accuracy and AUC, and
3.3. Temporal Analysis of Temporal Enhanced Ultrasound

Figure 3.9: Learning curves of different RNN cells using the optimum hyper-parameters in our search space. All of the models use the RMSprop optimizer and converge after 65±7 epochs.

The horizontal axis is the number of iterations. We observe that all models converge after 65±7 epochs, and GRU and LSTM cells outperform vanilla RNN cells in terms of accuracy. Comparing Fig. 3.9(a) and (b) demonstrates that the network with GRU cells has a steeper learning curve and converges faster than the network with LSTM cells. One possible reason could be the fewer number of parameters to be learned in GRU cells compared to LSTM cells. Fig. 3.9 shows that the network with LSTM cells leads to a lower loss value and a higher accuracy.

3.3.4.2 Model Performance

Table 3.4 shows the classification results in the test dataset, \( D_{test} \), including 171 target cores. Models with LSTM and GRU cells consistently achieve higher performance compared to vanilla RNN and the spectral method proposed in [11]. A two-way paired t-test shows statistically significant improvement in AUC \( (p < 0.05) \) with LSTM and GRU cells. Moreover, the LSTM configuration has the highest performance for detection of cancer. Using the LSTM model as the best configuration, we achieve specificity and sensitivity of 0.98 and 0.76, respectively, where we classify 31 out of 41 cancerous cores correctly. Table 3.5 and Table 3.6 shows performance of models for classification of cores in \( D_{test} \) for different MR suspicious levels as explained in Section 2.1. For samples with moderate MR suspicious level (70% of all cores), we achieve AUC of 0.97 using the LSTM-RNN structure. In this group, our sensitivity, specificity, and accuracy are 0.78, 0.98, and 0.95, respectively. For samples with high MR suspicious level, we consistently achieve higher sensitivity result compared to those with moderate MR suspicious level.
Table 3.4: Model performance for classification of cores in the test data (N = 171).

<table>
<thead>
<tr>
<th>Method</th>
<th>Specificity</th>
<th>Sensitivity</th>
<th>Accuracy</th>
<th>AUC</th>
</tr>
</thead>
<tbody>
<tr>
<td>LSTM</td>
<td>0.98</td>
<td>0.76</td>
<td>0.93</td>
<td>0.96</td>
</tr>
<tr>
<td>GRU</td>
<td>0.95</td>
<td>0.70</td>
<td>0.86</td>
<td>0.92</td>
</tr>
<tr>
<td>Vanilla RNN</td>
<td>0.72</td>
<td>0.69</td>
<td>0.75</td>
<td>0.76</td>
</tr>
<tr>
<td>Spectral [11]</td>
<td>0.73</td>
<td>0.63</td>
<td>0.78</td>
<td>0.76</td>
</tr>
</tbody>
</table>

Table 3.5: Model performance for classification of cores in the test data for Moderate MR suspicious level. N indicates the number of cores in each group.

<table>
<thead>
<tr>
<th>MR suspicious levels</th>
<th>Moderate MR suspicious level (N = 115)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Specificity</td>
</tr>
<tr>
<td>LSTM</td>
<td>0.98</td>
</tr>
<tr>
<td>GRU</td>
<td>0.95</td>
</tr>
<tr>
<td>Vanilla RNN</td>
<td>0.82</td>
</tr>
<tr>
<td>Spectral [11]</td>
<td>0.70</td>
</tr>
</tbody>
</table>

Table 3.6: Model performance for classification of cores in the test data for High MR suspicious level. N indicates the number of cores in each group.

<table>
<thead>
<tr>
<th>MR suspicious levels</th>
<th>High MR suspicious level (N = 20)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Specificity</td>
</tr>
<tr>
<td>LSTM</td>
<td>0.86</td>
</tr>
<tr>
<td>GRU</td>
<td>0.80</td>
</tr>
<tr>
<td>Vanilla RNN</td>
<td>0.85</td>
</tr>
<tr>
<td>Spectral [11]</td>
<td>0.83</td>
</tr>
</tbody>
</table>

3.3.4.3 Comparison with Other Methods

The best results to-date involving TeUS are based on spectral analysis as proposed in [11] and Section 3.2. We compare our RNN-based approach with this method as the most related work. For the fair comparison, we implement the spectral method with the same data division as proposed
for RNN-based method. As reported in Table 3.4, Table 3.5 and Table 3.6, the LSTM-RNN models consistently outperform spectral analysis of TeUS with a two-way paired t-test showing statistically significant improvement in sensitivity, specificity, accuracy, and AUC ($p < 0.05$) for LSTM-RNN. Datasets used in a few other studies are different from the current work so a detailed comparison is not feasible. However, the proposed network architectures outperform Imani et al. [74], who used cascade CNNs, and Uniyal et al. [153], who applied random forests for TeUS based prostate cancer detection, both on data from 14 patients. While analysis of a single RF ultrasound frame is not possible in the context of our current clinical study, previously, we have shown that TeUS is complementary to this information and generally outperforms analysis of single RF frames [36, 107]. The best results reported using a single RF frame analysis [42] involve 64 subjects with an AUC of 0.84, where they used PSA as an additional surrogate. In a recent study, Nahlawi et al. [115, 117] used Hidden Markov Models (HMMs) to model the temporal information of TeUS data for prostate cancer detection. In a limited clinical study including 14 subjects, they identified cancerous regions with an accuracy of 0.85. Furthermore, we qualitatively compare the cancer likelihood colormap resulting from LSTM-RNN with the spectral analysis of TeUS [11]. Figure 3.10 shows two examples of the cancer likelihood maps from the test dataset. There is an observable improvement in identifying cancerous regions around the biopsy target that match the gold-standard label.

### 3.3.4.4 Network Analysis

To analyze the network behavior and identify LSTM cells that contribute most to differentiating between benign and cancerous tissue, we examine the final high-level feature representation for cancerous and benign samples. By generating the difference map between the final activation of the network ($h_t$ at $t = 100$) for TeUS data from benign and cancerous samples, we identify 20 cells with the highest activation difference.

In these cells, we observe the evolution of high-level feature representations. Specifically, as per Algorithm 1, input TeUS data from benign and cancerous ROIs in $D_{test}^T$ are forward propagated in the models. Activation of the input gate ($i(t)$), the output gate ($o(t)$) and the cell state ($c(t)$) for the top 20 active cells are studied. We observe cell states $c(t)$ evolve over time and gradually learn discriminative information. Moreover, the input gate $i(t)$ evolves so that it attenuates parts of the input TeUS sequence and detects the important information from the sequence. Interestingly, the input gate
3.3. Temporal Analysis of Temporal Enhanced Ultrasound

![Figure 3.10: Cancer likelihood maps overlaid on B-mode ultrasound images, along with the projected needle path in the TeUS data, and centered on the target. Red indicates predicted labels as cancer, and blue indicates predicted benign regions. The boundary of the segmented prostate in MRI is overlaid on TRUS data. The arrow points to the target location. The top row shows the result of LSTM and the bottom row shows the result of spectral analysis [11] for benign targets (a), and cancer targets (b) and (c).](image)

reduces the contribution of TeUS data to model learning around time step 50, for both cancerous and benign samples. The evolution and attenuation patterns of \(c(t)\) and \(i(t)\) suggest that the most discriminative features for distinguishing cancerous and benign tissue samples are captured within the first half of TeUS sequence. These findings match those reported by Nahlawi et al. [115, 117].

To further examine this, we evaluate the evolving behavior of LSTM-RNN by training and testing the RNN models with different TeUS sequence lengths. Figure 3.11 shows the performance of the models evaluated by AUC for different TeUS sequence lengths. For each case, using the training procedure explained in Section 3.3.3.3 we trained an RNN-based deep network with 10-100 RNN cells corresponding to TeUS length. Similar to previous observations, the vanilla RNN-based model has the lowest performance compared to GRU and LSTM based models. By increasing the length of input TeUS sequence, the performance of the models improve. However, for TeUS sequence length more than 50, the improvement saturates. A two-way paired t-test shows that, for sequence length more than 50, there is no statistically significant
improvement in performance using the LSTM-RNN model ($p > 0.05$).

![Sequence Length effect: The most discriminative features for detection of prostate cancer can be learned from a fraction of the full TeUS time series.](image)

**Figure 3.11:** Sequence Length effect: The most discriminative features for detection of prostate cancer can be learned from a fraction of the full TeUS time series.

### 3.4 Conclusion

In this chapter, we focused on the development of probabilistic models for prostate cancer using temporal and spectral analysis of TeUS data. In the first section, we presented an approach for classification of tissue labels obtained in MR-TRUS guided targeted prostate biopsy using spectral analysis of TeUS. We utilized a DBN [13] for systematic learning of discriminant latent features from high-dimensional temporal-ultrasound-features for characterizing prostate cancer. We then applied an SVM classifier along with the activation of the trained DBN to characterize the prostate tissue. In a large clinical study including 255 TRUS-guided biopsy cores, we identified two important factors that affect our classification performance: i) distance of the target to the segmented prostate boundary which is correlated with the registration error between mp-MRI, TRUS, and histopathology, and ii) disagreement between the axial and sagittal histopathology results. In this section, we built our classification model using a fixed training dataset consisting of TeUS data of 32 biopsy cores and assessed the performance of the model on the remaining 223 biopsy cores. The test data is divided into three sub-groups according to the distance of the target to the prostate boundary and agreement between axial and sagittal histopathology labels. For cores from targets with moderate MR suspicious level in $D_{test}$, we achieved AUC...
of 0.80, where mp-MRI has low positive predictive value.

In this second part, we utilized deep RNNs to explicitly model the temporal information of TeUS for detecting prostate cancer. The investigation of several RNN structures showed that LSTM-based RNN can efficiently capture temporal patterns in TeUS data with statistically significant improvement in accuracy over our previously proposed spectral analysis approach [11]. We achieved AUC, sensitivity, specificity, and accuracy of 0.96, 0.76, 0.98, and 0.93, respectively. We also presented algorithms for in-depth analysis of high-level latent features of LSTM-based RNN and DBN. A transformational finding, achieved through this analysis, is that the most discriminative features for detection of prostate cancer can be learned from a fraction of the full TeUS time series. Specifically, in our data, less than 50 ultrasound frames were required to build models that accurately detect prostate cancer. This information can be used to optimize TeUS data acquisition for clinical translation. Moreover, our results showed that analysis of temporal ultrasound data is a promising technology for accurate classification of tissue labels that were identified in mp-MRI as suspicious and can potentially complement mp-MRI for TRUS-guided biopsy.
Chapter 4

Detection of High-Grade Prostate Cancer Using TeUS

_The world is noisy and messy. You need to deal with the noise and uncertainty._
— Daphne Koller

4.1 Introduction

Early diagnosis, and accurate grading and staging of prostate cancer play significant roles in the choice and the success of treatments [107]. Grading of prostate cancer is established by histopathological analysis of the obtained cores and staging determines the extent of the disease beyond the prostate itself. Therapeutic aspects of prostate cancer have progressed significantly, over the recent years, for patients with advanced disease [144]. However, men with indolent prostate cancer, who constitute the vast majority of diagnosed cases, are over-treated with the traditional options of surgery or...
4.1. Introduction

radiation therapy, leading to a decline in their quality of life. The utility of active surveillance for the effective disease management in men with indolent prostate cancer is dependent on accurate assessment of the disease grade and extent \[ 39, 143, 144 \]. Using a sensitive imaging modality for tissue-characterization and for biopsy guidance can substantially contribute to the appropriate and adequate treatment of prostate cancer. There have been a large number of efforts to adopt ultrasound-based tissue characterization for prostate cancer diagnosis during the biopsy procedure. Most contributions focus on the analysis of texture \[ 98 \] and spectral features \[ 42 \] within a single ultrasound frame. Elastography \[ 33 \] and Doppler imaging \[ 119 \] also aim to distinguish different tissue types based on their measured tissue stiffness and blood circulation, respectively. However, accurate characterization of aggressive from indolent prostate cancer and grading of the disease are still open issues \[ 82 \].

Over the past decade, features extracted from TeUS data have been used in a machine learning framework to predict labels provided by histopathology as the ground-truth. As we have also shown in Chapter 3, TeUS has been used successfully for characterization of cancerous and non-cancerous prostate tissue in \textit{ex vivo} \[ 107, 108 \] and \textit{in vivo} \[ 11, 13, 73, 75, 109 \] studies. In these studies, the area under the receiver operating characteristic curve (AUC) of 0.76-0.93 has been reported. TeUS has also been used to distinguish between various cancer grades in preliminary whole-mount studies \[ 82 \].

In this chapter, we focus on detection of higher grade prostate cancer and the problem of detection of different prostate cancer grades. From the machine point of view, this task can be modeled as a probabilistic multi-class classification task. 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. We try to address these challenges through two approaches. First, the learning of statistical distribution of cancer in the biopsy core using the automatically learned feature in spectral analysis of TeUS. Second, in a different approach, we embed the prior knowledge from the histopathology as the soft labels in a probabilistic model based on the analysis of temporal aspect of TeUS. This chapter is accordingly subdivided into two parts.
4.2 Prostate Cancer Grading Using Spectral Analysis of TeUS

In this section, we propose a cancer grading approach for transrectal ultrasound-guided prostate biopsy based on spectral analysis of TeUS signals. Histopathological grading of prostate cancer reports the statistics of cancer distribution in a biopsy core. The approach relies on a coarse-to-fine classification, similar to histopathology reporting, and uses statistical analysis and deep learning to determine the distribution of aggressive cancer in ultrasound image regions surrounding a biopsy target. Our approach consists of two steps; in the first step, we learn high-level latent features that maximally differentiate benign from cancerous tissue. In the second step, we model the statistical distribution of prostate cancer grades in the space of latent features. In a study with 197 biopsy cores from 132 subjects, our approach can effectively separate clinically significant disease from low-grade tumors and benign tissue. Further, we achieve the area under the curve of 0.8 for separating aggressive cancer from benign tissue in large tumors.

4.2.1 Materials

In this section, we use the data from our first study as we explained in Chapter 2 and we focus on the spectral representation of TeUS. Our findings from previous chapter indicate that the agreement between axial and sagittal histopathology of each target impact the classification accuracy in spectral analysis of TeUS. Thus, from the whole 255 cores in our dataset, we divide the data from 197 cores, who have the agreement between axial and sagittal pathology, into the train, $D^s_{\text{train}}$, and test, $D^s_{\text{test}}$, sets as explained in Table 4.1. For building the classification model, we need to use homogeneous tissue regions with the known gold standard. Given the potential mis-registration between MR and TRUS images, we use biopsy cores with at least 7 mm of cancer for a typical core length of 18 mm to build our model. Training data is made up of 32 biopsy cores from 27 patients with the following histopathology labels: 19 benign, 0 GS of 3+3, 5 GS of 3+4, 2 GS of 4+3, 4 GS of 4+4 and, 2 GS of 4+5. The test data consists of 165 cores from 114 patients, with the following distribution: 121 benign, 12 GS of 3+3, 14 GS of 3+4, 2 GS of 4+3, and 16 GS of 4+4.
4.2. Prostate Cancer Grading Using Spectral Analysis of TeUS

Table 4.1: Gleason score distribution in TeUS test and train dataset. Table represents the number of cores for each category.

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Benign</th>
<th>GS 3+3</th>
<th>GS 3+4</th>
<th>GS 4+3</th>
<th>GS 4+4</th>
<th>GS 4+5</th>
</tr>
</thead>
<tbody>
<tr>
<td>$D_{\text{train}}^S$</td>
<td>19</td>
<td>0</td>
<td>5</td>
<td>2</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>$D_{\text{test}}^S$</td>
<td>121</td>
<td>12</td>
<td>14</td>
<td>2</td>
<td>16</td>
<td>0</td>
</tr>
</tbody>
</table>

4.2.2 Method

Prostate cancer grading can be viewed as a multi-class classification problem where the objective is to detect the prostate cancer aggressiveness level (benign or Gleason grades 3, 4 or 5) for a given area in the tissue. Training a multi-class classifier given the histopathology reports of prostate biopsy as the ground-truth is non-trivial. Challenges associated with the histopathology labels for prostate grading are:

1. The ground-truth reports a measure of the statistical distribution of cancer in a biopsy core (e.g., % of cancer in core).
2. The exact location of the cancerous tissue in the core is not provided.
3. As a result of partial information as explained above, the exact label of each ROI in a core is not known.

Figure 4.8 shows an illustration of these ground-truth issues. The statistics of ROIs with various labels in a biopsy core are the only known information. In this approach, similar to the pathology reporting, we propose a coarse-to-fine classification approach to find a statistical representation of the distribution of ROIs in various classes (benign and Gleason grades 3, or 4). Figure 4.1 shows an illustration of the proposed grading method. There are two major steps: (1) feature learning to extract latent features that maximally separate benign from cancerous tissues; and (2) distribution learning to model the statistical distribution of prostate cancer grades in the space of learned features.

4.2.2.1 Feature Learning

As we explained in Chapter 3.2 we can use a DBN \cite{13, 22} to automatically learn a high-level latent feature representation of the TeUS data that can separate benign and cancerous lesions. In summary, the network structure
4.2. Prostate Cancer Grading Using Spectral Analysis of TeUS

Figure 4.1: An illustration of the proposed cancer grading approach using spectral analysis of TeUS.

includes 100, 50 and 6 hidden units in three consecutive layers, where the last hidden layer represents the latent features. In the pre-training step, the learning rate is fixed at 0.001, mini-batch size is 5, and the epoch is 100. Momentum and weight cost are set to defaults of 0.9, and $2 \times 10^{-4}$, respectively. In the discriminative fine-tuning step, we add a node to represent the labels of observations, and back-propagation with a learning rate of 0.01 for 70 epochs and mini-batch size of 10 is used. Due to a limited number of cancerous cores, we have not used any validation set in this work. In order to test the generalization of the trained DBN, we make certain that the test data is never used to pre-train or fine-tune the network parameters. The trained network maps the set of 50 spectral components for each ROI to six high-level latent features. Then, we perform dimensionality reduction in the space of the six latent features. We use Zero-phase Component Analysis [20] to whiten the features and to determine the top two eigen vectors, $f_1$ and $f_2$. We call this space the *eigen feature space.*
4.2. Prostate Cancer Grading Using Spectral Analysis of TeUS

4.2.2.2 Distribution Learning

To learn the distribution of different Gleason grades in the eigen feature space, we use a Gaussian Mixture Model (GMM) \[160\]. The \( K \)-component GMM is denoted by \( \Theta = \{ (\omega_k, \mu_k, \Sigma_k) | k = 1, ..., K \} \), where \( \omega_k \) is the mixing weight \( \sum_{k=1}^{K} \omega_k = 1 \), \( \mu_k \) is the mean and \( \Sigma_k \) is the covariance matrix of the \( k \)-th mixture component. Starting with an initial mixture model, the parameters of \( \Theta \) are estimated with Expectation-Maximization (EM) \[160\]. Since the EM algorithm is a local optimization method, it is particularly sensitive to the initialization of the model. Instead of random initialization, we use the prior knowledge from pathology to devise a simple but efficient method for this purpose.

Figure 4.2 shows an overview of the proposed GMM initialization technique. Let \( X_H \) be the set of all ROIs of the cores in training data with histopathology labels \( H \in \{ \text{benign}, \text{GS 3+4, GS 4+3, GS 4+4} \} \). First, we map the distribution of the ROIs from benign cores, \( X_{\text{benign}} \), in the eigen feature space; we observe two distinct clusters that span histopathology labels of normal and fibromuscular tissue, chronic inflammation, atrophy, and PIN. We use k-means clustering to separate the two clusters and consider the cluster with the maximum number of “normal tissue” ROIs as the dominant benign cluster, and the second cluster as a representative for other non-cancerous tissue. Next, we map ROIs in the training dataset that corresponds to the cores with GS 4+4, \( X_{\text{GS 4+4}} \) in the eigen feature space, to identify the dominant cluster that represents Gleason 4 pattern. Finally, we use all other ROIs from cancerous cores that correspond to GS 3+4 and GS 4+3 to determine the cluster for Gleason 3 pattern in the eigen feature space. We denote the centroid of all clusters by \( C = \{ C_{\text{benign}}, C_{\text{G4}}, C_{\text{G3}}, C_{\text{non-cancerous}} \} \). To initialize the \( K \)-component GMM, we set \( K = 4 \) to model the four tissue patterns with mean, \( \mu_k \), for each Gaussian component equal to the centroid of each cluster. We use the equal covariance matrices for all components and set \( \Sigma_k \) to the covariance of \( X_H \). Each \( \omega_k, k = 1, ..., K \) is randomly drawn from a uniform distribution between \( [0, 1] \) and normalized by \( \sum_{k=1}^{K} \omega_k \).

4.2.2.3 Prediction of Gleason Score

To predict the Gleason score of each core, we map the data from ROIs within that core to the eigen feature space (see Fig.4.1). Subsequently, we assign a label from \{ benign, G3, G4, non-cancerous \} to each ROI based on its proximity to the corresponding cluster center in the eigen feature space. To
4.2. Prostate Cancer Grading Using Spectral Analysis of TeUS

![Diagram]

Figure 4.2: An illustration of the proposed GMM initialization method.

determine a GS for a test core, \( Y \), we follow the histopathology guidelines where we use the ratio of the number of ROIs labeled as benign, G3 (\( N_{G3} \)) and G4 (\( N_{G4} \)) (e.g., a core with a large number of G4 and a small number of G3 ROIs has GS 4+3):

\[
Y = \begin{cases} 
\text{GS 4+3 or higher,} & N_{G4} \neq 0 \& N_{G4} \geq N_{G3} \\
\text{GS 3+4 or lower,} & N_{G3} \neq 0 \& N_{G4} < N_{G3} \\
\text{benign,} & \text{otherwise}
\end{cases}
\]

Since none of the cores labeled by histopathology as GS 3+3 did not make our selection criteria for training samples, we use cores with GS 3+4 and 4+3 to derive the features that describe Gleason grade 3.

4.2.3 Results and Discussion

To assess the performance of our method, we use sensitivity, specificity, and accuracy of the approach in detecting cancerous tissue samples. We consider all cancerous cores as the positive class and other non-cancerous cores as the negative class. Sensitivity or recall is the percentage of cancerous cores that are correctly identified while specificity is the proportion of non-cancerous cores that are correctly classified. Accuracy is the ratio of the true results (both true positives and true negatives) over the total number of cores. We also report the overall performance of our approach using the area under the receiver operating characteristic curve (AUC). Receiver Operating Characteristic (ROC) is a two-dimensional graph of sensitivity
4.2. Prostate Cancer Grading Using Spectral Analysis of TeUS

Table 4.2: Model performance for prostate cancer grading in the test dataset using TeUS only and by integration of TeUS and mp-MRI. \( L \) is the largest length of the tumor visible in mp-MRI.

<table>
<thead>
<tr>
<th>Method</th>
<th>TeUS All cores</th>
<th>TeUS + mp-MRI All cores</th>
<th>TeUS All cores</th>
<th>TeUS + mp-MRI All cores</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non-cancerous vs. GS ≥ 4+3</td>
<td>0.69</td>
<td>0.80</td>
<td>0.82</td>
<td>0.89</td>
</tr>
<tr>
<td>Non-cancerous vs. GS ≤ 3+4</td>
<td>0.62</td>
<td>0.63</td>
<td>0.67</td>
<td>0.81</td>
</tr>
<tr>
<td>GS ≤ 3+4 vs. GS ≥ 4+3</td>
<td>0.61</td>
<td>0.67</td>
<td>0.71</td>
<td>0.71</td>
</tr>
</tbody>
</table>

versus (1-specificity), depicting the relative trade-offs between sensitivity and specificity. Accuracy can be reported at any particular threshold on this curve. The maximum AUC is 1, where larger AUC values indicate better classification performance [40].

4.2.3.1 Prostate Cancer Detection and Grading

Tables 4.2 and 4.3 show the classification and grading performance based on the inter-class AUC, accuracy, sensitivity, and specificity. To investigate the effect of the size of the tumor on our detection performance, we also show the AUC versus the largest length of the tumor in MRI. This length ranges from 0.3 cm to 3.8 cm in our dataset. As seen in Table 4.2, in cores with MR-tumor-size ≥ 2.0 cm, we obtain AUC of 0.80 in classifying cores with GS ≥ 4+3 from non-cancerous cores. Moreover, we achieve AUC, accuracy, sensitivity, and specificity of 0.70, 70%, 70%, and 71%, respectively, in the detection of cancerous from non-cancerous cores with MR-tumor-size ≥ 2.0 cm.

Figure 4.3 shows examples of the cancer likelihood maps from test dataset, derived from the output of the proposed clustering algorithm. Cancer colormaps overlaid on B-mode ultrasound image, along with the projected needle path in the TeUS data and centered on the target. The red boundary shows the segmented prostate in MRI projected in TRUS coordinates. In the colormaps, red and yellow regions show 0.5 × 0.5 mm × mm ROIs which we detect as Gleason grades 4 and 3, respectively. The blue areas belong to non-cancerous ROIs.

Figure 4.3(b) shows a case where the axial and sagittal histopathology results do not agree (see Section 3.2). The axial pathology indicates GS 3+4 whereas the sagittal pathology reports the core as GS 4+4. The colormap demonstrates using our approach, the clinician could have reoriented the
4.2. Prostate Cancer Grading Using Spectral Analysis of TeUS

(a) MRI lesion length = 27 mm, benign target
(b) MRI lesion length = 36 mm, GS ≤ 3 + 4
(c) MRI lesion length = 24 mm, GS ≤ 3 + 4
(d) MRI lesion length = 17 mm, GS ≥ 4 + 3

Figure 4.3: Cancer likelihood maps overlaid on B-mode US image, along with the projected needle path in the TeUS data and centered on the target. The ROIs for which we detect as Gleason grade of 4 and 3 are colored in red and yellow, respectively. The non-cancerous ROIs are colored as blue. The red boundary shows the segmented prostate in MRI projected in TRUS coordinates and the arrow pointer shows the target.

needle to biopsy a more aggressive region.

4.2.3.2 Integration of TeUS and mp-MRI

To take advantage of mp-MRI information, we combine the TeUS cancer detection results with readings from mp-MRI suspicious levels. If mp-MRI declares cancer suspicious level as low or high for a core, we use that prediction alone and report the core as benign or aggressive cancer, respectively. However, when mp-MRI declares the suspicious level as intermediate (70% of all cores in our data), we use predictions based on TeUS. For tumors with \( L \geq 2.0 \text{ cm} \), the integrated approach leads to an AUC of 0.76 for predicting cancer grades compared that of 0.65 using mp-MRI alone, and 0.70 using TeUS data alone. Moreover, for classification of cores with GS ≥ 4+3 from non-cancerous cores, the combined AUC is 0.82. The results indicate both
4.2. Prostate Cancer Grading Using Spectral Analysis of TeUS

Table 4.3: Model performance for classification of cancerous vs. non-cancerous cores in the test dataset using TeUS only and Integration of TeUS and mp-MRI. \( L \) is the greatest length of the tumor visible in mp-MRI.

<table>
<thead>
<tr>
<th>Evaluation</th>
<th>TeUS All cores</th>
<th>( L \geq 2.0 \text{ cm} )</th>
<th>TeUS + mp-MRI All cores</th>
<th>( L \geq 2.0 \text{ cm} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accuracy</td>
<td>65%</td>
<td>70%</td>
<td>68%</td>
<td>70%</td>
</tr>
<tr>
<td>Sensitivity</td>
<td>62%</td>
<td>70%</td>
<td>63%</td>
<td>70%</td>
</tr>
<tr>
<td>Specificity</td>
<td>67%</td>
<td>71%</td>
<td>67%</td>
<td>72%</td>
</tr>
<tr>
<td>AUC</td>
<td>0.65</td>
<td>0.70</td>
<td>0.72</td>
<td>0.76</td>
</tr>
</tbody>
</table>

TeUS and integration of TeUS with mp-MRI have higher performance for larger tumors.

4.2.3.3 Sensitivity Analysis

Performance in Anatomical Zones: Figure 4.4 (top) summarizes the biopsy target locations, distribution, and histopathology outcomes for the test data. The prostate region is divided into anterior/posterior, and central/peripheral zones for the base, midgland, and apex regions. In our test data, 34% (19 out of 56 biopsies) of all cancerous cores were in the central region where 24% (25 out of 109 biopsies) were in the peripheral region. Figure 4.4 (middle) shows the predictions of prostate cancer grades using TeUS in different anatomical zones. Although more biopsies were performed in the peripheral zone, a higher portion of positive biopsies was observed in the central zone. In the central zone, we can differentiate between non-cancerous targets and clinically significant cancer (GS \( \geq 4 + 3 \)) with an AUC of 0.80. The bottom row depicts prostate cancer grading performance integrating TeUS and MRI information.

Choice of the Training Data: We also analyzed the sensitivity of our approach to the choice of training data. To create new training and test sets we permute our original data where, in each permutation, we exchange a randomly selected cancerous or benign core in the primary training and test data. This results in 32 different data divisions. As Fig. 4.5 shows, the average AUC of the sensitivity analysis follows our previous performance results, which supports the generalization of the proposed model.

Size of the Training Data: We also investigate the effect of training data
4.2. Prostate Cancer Grading Using Spectral Analysis of TeUS

Figure 4.4: Target location and distribution of biopsies in the test data. Light and dark gray indicate central and peripheral zones, respectively. The pie charts show the number of cores and their histopathology. The size of the chart is proportional to the number of biopsies (in the range from 1 to 25), and the colors dark red, light red and blue refer to cores with GS $\geq 4 + 3$, GS $\leq 3 + 4$ and benign pathology, respectively. The top, middle, and bottom rows depict histopathology results, TeUS prediction, and integration of TeUS and MRI, respectively.

size on the grading performance by gradually increasing the size of the dataset from 16 to 56 cores. When choosing new training samples to add, we use biopsy cores with at least 4.0 mm of cancer for a typical core length of 18 mm (given the potential mis-registration between MR and TRUS images [11] it is a prudent step). Otherwise, we randomly select an equal number of benign
4.2. Prostate Cancer Grading Using Spectral Analysis of TeUS

Figure 4.5: Model performance for prostate cancer grading using spectral analysis of TeUS and distribution learning in the test dataset and permutation set.

Figure 4.6: Model performance for different sizes of training dataset using spectral analysis of TeUS and distribution learning.

and cancerous cores. Figure 4.6 shows the AUC, of differentiation between prostate cancer and non-cancerous tissue, versus the size of the dataset. In general, there is an increasing trend for the AUC and our approach has a higher performance using more training samples.

**Number of Features:** It has been shown that two features ($f_1$ and $f_2$) are very effective in classification. We further analyze the effect of including any of the six features in the model over the final results. Figure 10 shows the
4.3 Temporal Analysis of TeUS for prostate cancer grading

Despite promising results in detecting high grade prostate cancer using spectral analysis of TeUS and distribution learning, accurate characterization of aggressive lesions from indolent ones still requires refinement. We need to have a more precise grading methods based on temporal analysis of TeUS which can enable real-time depiction of cancer likelihood map during the biopsy procedure. In addition, as we discussed in the past sections, the goodness of models built based on all the above analyses depends on detailed, noise-free annotations of ground-truth labels from pathology. Histopathology data used for training of the computer-aided diagnosis models are sparsely
4.3. Temporal Analysis of TeUS for prostate cancer grading

Figure 4.8: 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.

annotated with the inevitable ubiquity of noise.

As we explained in previous sections, histopathology reports include the length of cancer in the biopsy core and a Gleason Score (GS) [119]. 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 tissue) to 5 (aggressive cancerous tissue). 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. 4.8). 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 this section, we propose a method to directly alleviate the challenge of sparse and noisy histopathology ground-truth labels to improve TeUS-based prostate biopsy guidance. 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
4.3. Temporal Analysis of TeUS for prostate cancer grading

of 0.84 for cancer grading and mean squared error of 0.12 in the estimation of tumor in biopsy core length.

4.3.1 Method

4.3.1.1 Discriminative Model

Let \( \mathcal{D}^T = \{(x^{(i)}, y^{(i)})\}_{i=1}^{|\mathcal{D}^T|} \) represent the collection of all ROIs, where \( x^{(i)} \) is the \( i \)th TeUS sequence with length \( T \) and is labeled as \( y^{(i)} \) 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 \( \mathcal{D}^T \). 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 [47]. 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 (Fig. 4.9). In the heart 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 input sequence \( x = (x_1, ..., x_T) \), RNN computes the hidden vector sequence \( h = (h_1, ..., h_T) \) in the sequence learning step. This hidden vector, \( h \) is a function of the input sequence \( x \), model parameters, \( \Theta \), and time.

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

\[
\overline{y}_b^{(i)} = \arg \max_j S(z_j^{(i)}), \quad j \in \{0, 1\}, \quad z^{(i)} = W_s^T h + b_s, \quad (4.1)
\]

where \( W_s \) and \( b_s \) are the weight and bias of the fully-connected layer and \( S \) 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
4.3. Temporal Analysis of TeUS for prostate cancer grading

Figure 4.9: Overview of the second stage in the proposed method: the goal of this stage is to assign a pathological score to a sample. To mitigate the problem of imperfect and noisy labels, we embed the length of cancer in the ground-truth probability vector as a soft label.

$$y^{(i)}_b$$ and $${\overline y}^{(i)}_b$$ over all training samples.

**Stage 2: Grading of Cancerous Samples:** The goal of this stage is to assign a pathological score to $${\mathcal D}_{cancer}^{train} = \{(x^{(i)}, y^{(i)}_b) \in {\mathcal D}_{train}^T \mid y^{(i)}_b = 1\}_{i=1}^N$$. Here, unlike the first stage, we are facing a multi-label classification task with sparse labeling (Fig. 4.8). The histopathology reports include two informative parts:

1. Gleason score which implies any of the possible labels for all ROIs within a core, $${\Omega} \in \{\text{Benign, G3, G4}\}$$. In this representation, all or at least two of these patterns can happen at the same time (Fig. 4.8). As we explained before, we can also interpret a given Gleason score as a measure distribution of each grade in a core.

2. In our dataset, we also have access to the measured length of cancerous tissues ($$\text{Len}$$) in a typical core length ($$\text{Len}_{\text{typical}}$$) 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.
4.3. Temporal Analysis of TeUS for prostate cancer grading

methods. For this purpose, using $D_{cancer}^{train}$ the output of sequence learning step $h$ is fed into a $k$-way softmax function, which produces a probability distribution over the $k$ possible class labels ($k = 3$). Suppose $Len(i)$ represents the length of cancer for the core that $x(i)$ belongs to. The ground-truth probability vector of the $i^{th}$ ROI is defined as $p^{(i)} = [p_{1}^{(i)}, ..., p_{k}^{(i)}]$. To estimate these probabilities we define the normalized cancer percentage as $C^{(i)} = \frac{Len(i)}{Len_{typical}} \in [0, 1]$. For $k = 3$:

$$p^{(i)} = \left[ p_{1}^{(i)} = (1 - C^{(i)}), p_{2}^{(i)} = \omega \times C^{(i)}, p_{3}^{(i)} = (1 - \omega) \times C^{(i)} \right],$$  \hspace{1cm} (4.2)

where $\omega$ 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, $\omega$ 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 GS 4+4 readings, $\omega$ is set to 1 and 0, respectively. Then, the cost function to be minimized is defined as:

$$J = \frac{1}{|D_{cancer}^{train}|} \sum_{i=1}^{N} \sum_{k=1}^{K} (p_{k}^{(i)} - \bar{p}_{k})^2,$$

where $\bar{p}^{(i)} = [\bar{p}_{1}^{(i)}, ..., \bar{p}_{k}^{(i)}]$ is the predictive probability vector.

4.3.1.2 Cancer Grading and Tumor in Core Length Estimation

Suppose $C^{T} = \{ (x^{(i)}, y^{(i)}) \}_{i=1}^{|C^{T}|}$ represent the collection of all labeled ROIs surrounding a target core, where $C^{T} \in D_{test}^{T}$, $|C^{T}| = 80$, $x^{(i)}$ represents the $i^{th}$ TeUS sequence of the core, and $y^{(i)}$ 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, $\bar{y}^{(i)}$, as 1, when $P(y^{(i)} \mid x^{(i)}) \geq 0.5$, and as 0 otherwise. The probability of a given core being cancerous based on the cancerous ROIs within that core is $P_b = \sum_{i=1}^{|C^{T}|} I(\bar{y}^{(i)} = 1) / |C^{T}|$. A binary label of 1 is assigned to a core, when $P_b \geq 0.5$. 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 $p_{m}^{(i)} = [p_{1}^{(i)}, p_{2}^{(i)}, p_{3}^{(i)}]$ represents the predictive probability output of $i^{th}$ TeUS
sequence in the second stage. We define the average predictive probability as:

\[ P_m = \frac{1}{|C'^T|} \sum_{i=1}^{\|C'^T\|} p_m(i), \]  

(4.4)

Following the histopathology guidelines, to determine a GS for a cancerous test core, \( Y \), we define the core as “GS 4+3 or higher” when \( P_m^{(3)} \geq P_m^{(2)} \) and otherwise as “GS 3+4 or lower”. Furthermore, based on Equation (4.4) and (4.2) we can estimate the predicted length of cancer for this core as:

\[ \text{Len}^C = (1 - P_m^{(1)}) \times \text{Len}^\text{typical}, \]  

(4.5)

### 4.3.1.3 Model Uncertainty Estimation

We also aim to estimate the model uncertainty in the detection of cancer for the areas outside the cancer foci, where the annotation is not available. The key for estimating model uncertainty is the posterior distribution \( P(\Theta|D) \), also referred to as Bayesian inference [48]. Here, we follow the idea in [48] to approximate model uncertainty using Monte Carlo dropout (MC dropout). Given a new input \( x^{(i)} \), we compute the model output with stochastic dropouts at each layer. That is, randomly drop out each hidden unit with certain probability \( p \). This procedure is repeated \( B \) times, and we obtain \( \{y_b^{*(1)}, ..., y_b^{*(B)}\} \). Then, the model uncertainty can be approximated by the sample variance:

\[ \frac{1}{B} \sum_{j=1}^{B} (\hat{y}_b^{*(j)} - \check{y}_b^{*(j)})^2, \]  

(4.6)

where \( \check{y}_b^{*(j)} \) is the average of \( y_b^{*(j)} \) values.

### 4.3.2 Experiments and Results

#### 4.3.2.1 Data Division and Model Selection

Data is divided into mutually exclusive patient sets for training, \( D_{train}^T \), and test, \( D_{test}^T \). 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 Section [2.3] we obtain a total...
Table 4.4: Model performance for classification of cores in the test data (N = 170). AUC$_1$, AUC$_2$ and AUC$_3$ refer to detection of Benign vs. GS$\leq 3+4$, Benign vs. GS$\geq 4+3$, and GS$\leq 3+4$ vs. GS$\geq 4+3$, respectively.

<table>
<thead>
<tr>
<th>Method</th>
<th>AUC$_1$</th>
<th>AUC$_2$</th>
<th>AUC$_3$</th>
<th>Average AUC</th>
</tr>
</thead>
<tbody>
<tr>
<td>LSTM</td>
<td>0.96</td>
<td>0.96</td>
<td>0.86</td>
<td>0.93</td>
</tr>
<tr>
<td>GRU</td>
<td>0.92</td>
<td>0.92</td>
<td>0.84</td>
<td>0.89</td>
</tr>
<tr>
<td>Vanilla RNN</td>
<td>0.76</td>
<td>0.76</td>
<td>0.70</td>
<td>0.74</td>
</tr>
<tr>
<td>BL-1</td>
<td>0.96</td>
<td>0.96</td>
<td>0.68</td>
<td>0.86</td>
</tr>
<tr>
<td>BL-2</td>
<td>0.75</td>
<td>0.68</td>
<td>0.58</td>
<td>0.67</td>
</tr>
<tr>
<td>BL-3</td>
<td>0.82</td>
<td>0.84</td>
<td>0.65</td>
<td>0.77</td>
</tr>
<tr>
<td>LSTM + GMM-Clustering</td>
<td>0.60</td>
<td>0.74</td>
<td>0.69</td>
<td>0.68</td>
</tr>
<tr>
<td>DBN + GMM-Clustering</td>
<td>0.68</td>
<td>0.62</td>
<td>0.60</td>
<td>0.63</td>
</tr>
</tbody>
</table>

number of $80 \times 1,536 = 122,880$ training samples ($N = |D_{\text{train}}| = 122,880$).

We use 20% of $D_{\text{train}}$ data as the held-out validation sets ($D_{\text{val}}$) 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, $\omega$, with three different optimization algorithms, SGD, RMSprop and Adam. Results from hyper-parameter 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, $d_r = 0.2, l_{\text{reg}} = 0.0001$ generate the lowest loss and the highest accuracy in $D_{\text{val}}$. 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 $D_{\text{train}}$ for training a model for the first stage and $D_{\text{train}}^{\text{cancer}}$ for the second stage model.

4.3.2.2 Comparative Method and Baselines

We use standard evaluation metrics as prior approaches [10, 11] to quantify our results. We assess the inter-class area under the receiver operating characteristic curve (AUC) for detection of Benign vs. GS$\leq 3+4$ (AUC$_1$), Benign vs. GS$\geq 4+3$ (AUC$_2$), and GS$\leq 3+4$ vs. GS$\geq 4+3$ (AUC$_3$). Table 4.4 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 $p^{(i)}$ with the labels from multi-label binarization as shown in Fig. 4.8 (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], and for GS 4+3 the encoded vector is [0, 0.3, 0.7]. We have also implemented the GMM-clustering method proposed in previous section for the current data division [11]. We have used the learned feature vector from Deep Belied Network (DBN) method [11] 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≤3+4 from GS≥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 [47].

4.3.2.3 Tumor in Core Length Estimation

Figure 4.10 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.

4.3.2.4 Cancer Likelihood Colormaps

Figure 4.11(a) shows an example of a cancer likelihood map for biopsy guidance derived from the output of the proposed two-stage approach. Figure 4.11(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.
4.4 Conclusion

Determining the aggressiveness of prostate cancer can help reduce the current high rate of over-treatment in patients with indolent cancer. In this chapter, we proposed methods for detection of higher grade prostate cancer using spectral and temporal analysis of TeUS data. In the first part, using the data from an \textit{in vivo} TRUS-guided prostate biopsy study, temporal enhanced ultrasound data was used to differentiate between clinically less significant prostate cancer (GS$\leq 3+4$), aggressive prostate (GS$\geq 4+3$) and non-cancerous prostate tissues. We utilized an unsupervised distribution learning approach to address the challenges related to ground-truth labeling in prostate cancer grading. First, we learned the differentiating features for detection of cancerous and non-cancerous prostate tissue, and then the statistical distribution of prostate cancer grades was modeled using a GMM. We showed that we could successfully differentiate among aggressive prostate cancer (GS$\geq 4+3$), clinically less significant prostate cancer (GS$\leq 3+4$), and non-cancerous prostate tissues. Furthermore, the combination of temporal enhanced ultrasound and mp-MRI had shown the potential to outperform either modality alone in the detection of prostate cancer. An AUC of 0.8 was achieved for separation of aggressive prostate cancer from non-cancerous tissue for tumors that were larger than 2.0 cm in their greatest dimension in MRI. Additionally, suspicious levels from MRI, when added to TeUS information, led to an AUC of 0.89 for classification of aggressive prostate cancer from non-cancerous tissue for tumors larger than 2.0 cm in MRI.

Figure 4.10: Scatter plot of the reported tumor in core length in histopathology vs. the predicted tumor in core length.
4.4. Conclusion

Figure 4.11: (a) Cancer likelihood maps overlaid on B-mode US image, along with the projected needle path in the TeUS data (GS \(\geq 4 + 3\)) and centered on the target. The ROIs of size 0.5 \(\times\) 0.5 \(mm\) \(\times\) \(mm\) 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 and the arrow pointer shows the target.[blue=low uncertainty, red=high uncertainty]

In the second part, we tried to address 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 suggested 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.
Chapter 5

Decision Support System for Prostate Biopsy Guidance

You’re either part of the solution or you’re part of the problem.
— Eldridge Cleaver

5.1 Introduction

Conventional ultrasound images, referred to as Brightness-mode (B-mode), are generated following envelope-detection of back-scattered radio frequency (RF) signals from tissue. B-mode gray-scale images are subjected to various non-linear processing steps, which are used to improve the visualization of the displayed image for assessment by physicians [141]. RF signals contain information about tissue microstructure at spatial scales much smaller than the conventional B-mode imaging resolution [126, 141]. Therefore, from a data processing perspective, using back-scattered RF signals with richer information content for tissue characterization can positively affect diagnostic decisions [36, 42, 111, 126]. We have shown that, TeUS RF can be used effectively for characterization of prostate cancer. In the \textit{in vivo} studies discussed in Chapter 3 and Chapter 4, we have achieved the areas under receiver operating characteristic curve (AUC) of 0.8-0.96 for grading and

---

diagnosis of prostate cancer \cite{73, 75} using spectral and temporal analysis of TeUS. Despite the promising results of TeUS for prostate cancer diagnosis, development and deployment of TeUS as well as all other ultrasound RF-based methods for computer-aided diagnosis deal with a common challenge: RF data is not available on all commercial ultrasound scanners and is usually only provided for research purposes. To address this challenge, we aim to identify an intermediate method that enables the use of B-mode data from conventional scanners for tissue characterization with TeUS.

B-mode images are a result of multiple non-linear processing steps applied to ultrasound RF signal including demodulation, non-linear intensity mapping, compression and filtering, which are specific to every scanner and are usually not disclosed. The lack of detailed information about the processing pipeline makes RF reconstruction from B-mode data almost impossible. Previous research has focused on the estimation of the parameters of compression and other non-linear operations to revert B-mode image generation and reconstruct RF data \cite{141}. However, this approach cannot be easily adopted for clinical applications as results need to be generated in real-time for different scanner types and various settings \cite{141}.

To overcome the challenge of accessibility of RF data, in the TeUS machine learning framework, we propose to use a transfer learning method \cite{130} to transfer knowledge between TeUS RF and B-mode data within an ultrasound scanner. Transfer learning is a method that can compensate for the diversity between datasets by learning from a source dataset (RF time series data) and apply the knowledge to a target dataset (B-mode time series data). This approach exploits the common information between the two domains. Common elements between two data domains can be features of the data or parameters of the models that are built using the data \cite{130}. Transfer learning has been applied to integrate data from multiple centers for characterization of plaque in carotid arteries \cite{155}, for protocol-invariant models for segmentation of brain MRI images \cite{154}, and for MRI data harmonization \cite{105}. Transfer learning combined with deep learning has shown tremendous potential in computer vision as well \cite{134}. In case of limited training examples, one can leverage pre-trained deep networks from alternate large scale data (e.g., ImageNet \cite{87}) in a transfer learning framework to extract features and employ them for the task at hand. Transfer learning has also been used for computer-aided detection, where information from images of natural scenes has been successfully transferred for detection of lung lesions \cite{142}. Deep learning methods, such as Deep Belief Networks (DBN) and Recurrent Neural Networks (RNNs), have an essential characteristic that makes them well suited for transfer learning: they can identify abstract features that
5.2 Transfer Learning From TeUS RF to B-mode Spectral Features

In this section, in a feasibility study, we present a method for prostate cancer detection using ultrasound TeUS data obtained either from Radio Frequency (RF) ultrasound signals or B-mode images. We demonstrate that by applying domain adaptation and transfer learning methods, a tissue classification model trained on TeUS RF data (source domain) can be deployed for classification using TeUS B-mode data alone (target domain), where both data are obtained on the same ultrasound scanner. This is a critical step for clinical translation of tissue classification techniques that primarily rely on accessing RF data since this imaging modality is not readily available on all commercial scanners in clinics. The proof-of-concept is provided for in vivo characterization of prostate cancer using TeUS B-mode data, where different non-linear processing filters in the pipeline of the RF to B-mode conversion result in a distribution shift between the two domains. Our in vivo study includes data obtained in MRI-guided targeted procedure for prostate biopsy from 84 subjects. We achieve a comparable area under the curve using TeUS
5.2. Transfer Learning From TeUS RF to B-mode Spectral Features

RF and B-mode data for medium to large cancer tumor sizes in biopsy cores (> 4 mm). Our result suggests that the proposed adaptation technique is successful in reducing the divergence between TeUS RF and B-mode data.

5.2.1 Materials

In this section, we use both annotated TeUS data from regions around the target location and also un-annotated TeUS data from the whole biopsy scan and inside the prostate boundary in ultrasound scan.

5.2.1.1 Unlabeled Data

As we mentioned in Section 2.3.2 in spectral representation of TeUS, each ROI is represented by \( F = 50 \) spectral features over the positive frequency range of 0 – 10 Hz. Unlike our typical biopsy region analysis, here, the RF spectral features from the ROIs of the entire imaging plane of the prostate and in scan level make up the unlabeled source dataset, \( D_{US}^S \), while the corresponding B-mode spectral features make up the unlabeled target dataset, \( D_{UT}^S \). Both \( D_{US}^S \) and \( D_{UT}^S \) have an equal number of samples. In total, \( D_U^S \) includes 5,255,680 unlabeled training samples, consisting of 2,627,840 RF and 2,627,840 B-mode time series data.

5.2.1.2 Labeled Data

For building the classification model, we randomly select an equal number of benign and cancerous cores, from the samples with at least 4.0 mm of cancer for a typical core length of 18 mm. Our labeled dataset consists of 84 biopsy cores with the following histopathology label distribution: Benign: 42 cores; GS 3+3: 2; GS 3+4: 14; GS 4+3: 3; GS 4+4: 18; and, GS 4+5: 5 cores. Similar to the unlabeled dataset, we extract spectral features from RF and B-mode time series of each ROI. The labeled source dataset, \( D_{LS}^S \), includes RF spectral features of ROIs selected from the above 84 cores and the target dataset \( D_{LT}^S \), includes the corresponding B-mode spectral features, where \( D_L^S = D_{LT}^S \cup D_{LS}^S \). Both \( D_{LS}^S \) and \( D_{LT}^S \) datasets have an equal number of samples of 6,720.

5.2.2 Methods

As discussed above, RF data is not readily accessible in all commercial ultrasound scanners. To potentially enable large-scale clinical deployment of TeUS-based tissue typing, we propose to use B-mode time series data (instead
5.2. Transfer Learning From TeUS RF to B-mode Spectral Features

Figure 5.1: An illustration of the proposed approach for domain adaptation between RF and B-mode time series data in TeUS framework.

of RF time series) along with a domain adaptation method. Figure 5.1 shows a block diagram of the solution. Our approach consists of two major steps:

1. Unsupervised learning of a common latent feature space using $D^S_U$ in a DBN training framework for the spectral analysis of TeUS.

2. Supervised learning of a shared support vector machine (SVM) classifier using the labeled dataset, $D^S_L$ for tissue typing.

We will discuss in detail these two steps and the intuition behind them in the following sections.
### 5.2.2.1 Unsupervised Domain Adaptation

In the proposed transfer learning setting, both the source (RF time series) and the target (B-mode time series) data are represented as spectral features. However, the distributions of the data in these two domains are different. We assume that, due to the intrinsic similarity of RF and B-mode time series domains, it is possible to derive a common feature representation to minimize domain divergence.

For this purpose, we start our unsupervised domain adaptation with a layer of Principal Component Analysis (PCA) to whiten the spectral features and align the source and target PCA subspaces. PCA provides a projection of each input vector to a low-dimensional coordinate vector, implicitly defining a low-dimensional hyperplane in input space near which information density is likely to be concentrated. We keep $d = 11$ top eigen vectors that retrain 99% of the variance of both TeUS RF and B-mode data in $D_{US}^S \subset \mathbb{R}^{F=50}$. Since, even after PCA projection, we keep a fairly large number of eigen directions, the main effect of this step is to smooth the input shared distribution by eliminating some of the variations involved in the least globally varying directions. Furthermore, by using whitening along with PCA projection, features are normalized to a unit variance [21] which is a key step in a successful learning process. Following the PCA whitening step, we train a shared DBN using both source and target data in the PCA subspace to minimize the domain divergence. The proposed DBN network structure includes two cascaded Restricted Boltzmann Machines (RBM) with $d = 11$ visible input units and the same number of hidden units in the last layer which represent the latent common features space (see Figure 5.1).

Given the two domains, $D_{US}^S$ and $D_{UT}^S$, there are two factors to be taken into consideration for unsupervised common representation learning between the source and target data:

- The first is to minimize the reconstruction error for both source and target domain data, as we explained in Section 3.2.1.2.

- The second is to minimize the divergence between the source and target distributions. For this purpose as recommended in literature [21], we use the Kullback–Leibler (KL) divergence as the measure of divergence.

We start with unsupervised pre-training of DBN with the greedy layer-wise procedure where the goal is to minimize the reconstruction error for the input data, and each layer is trained as an RBM by contrastive divergence [62]. In the next step, we fine-tune the DBN using Stochastic Gradient Descent (SGD), where the goal is to minimize:
5.2. Transfer Learning From TeUS RF to B-mode Spectral Features

- The cross-entropy between $\mathcal{D}_{UT}$ and the output of the shared network for $\mathcal{D}_{US}$.

- The cross-entropy between $\mathcal{D}_{US}$ and the output of the network for $\mathcal{D}_{UT}$.

Let $\tilde{f}_S^{(i)}$ indicates the $i^{th}$ RF TeUS sample in $\mathcal{D}_{US}$ and $\tilde{f}_T^{(i)}$ is the corresponding $i^{th}$ B-mode TeUS sample in $\mathcal{D}_{UT}$ after PCA projection. Also, $\bar{o}_S^{(i)}$ and $\bar{o}_T^{(i)}$ indicate the output of the DBN for $\tilde{f}_S^{(i)}$ and $\tilde{f}_T^{(i)}$, respectively. The size of the visible and the last hidden unit in our DBN is set to be the same and equal to $d = 11$. Thus, all $\tilde{f}_S^{(i)}$, $\tilde{f}_T^{(i)}$, $\bar{o}_S^{(i)}$, and $\bar{o}_T^{(i)}$ are $\subset \mathbb{R}^{d=11}$. Based on the Equation 3.7 in Section 3.2.1.2, the fine-tuning stage minimizes the above explained cross-entropy errors as the loss function, $\mathcal{L}_{Div}$:

$$
\mathcal{L}_{Div} = - \sum_{i \in \mathcal{D}_{US}} \tilde{f}_T^{(i)} \log \bar{o}_S^{(i)} - \sum_{i \in \mathcal{D}_{UT}} \tilde{f}_S^{(i)} \log \bar{o}_T^{(i)} ,
$$

(5.1)

Using the proposed network structure, we could use target samples as the ground-truth for the source sample and vice versa. The intuition behind our proposed approach is that a loss function defined by cross-entropy is closely related to the Kullback–Leibler (KL) divergence between the source and target domain [21], which is widely used to solve domain adaptation problems. KL divergence or relative entropy is a non-symmetric measure of divergence between two probability distributions, where minimizing the KL divergence of the network output for the source and target instances aims to draw the two distributions of the source and target domains similar in the common feature space. Minimizing the symmetrical version of KL divergence between the source and target distribution is equal to minimizing cross-entropy between the source and target distributions in the common feature space and vice versa [163]. By minimizing cross-entropy between the source and target instances in the input and the output latent-feature spaces, we modify the optimization problem to learn a model that can be better generalized for both source and target with less diversity between the two domains.

This approach can be simplified, by just considering the model that have to work with only, target data (here TeUS B-mode data). In this case, instead of minimizing the symmetrical version of KL divergence, the goal is to reduce the cross-entropy only between the target domain as the input of the network and the source domain as the preferred output of the network. Thus, in the plain version of the approach, the loss function in Equation 5.1 is only contained the second term.
5.2. Transfer Learning From TeUS RF to B-mode Spectral Features

5.2.2.2 Supervised Classification

After unsupervised training of the deep network, we use the labeled dataset, $D^S_L$, for supervised training of a shared classifier to separate cancerous from benign ROIs. We use the trained DBN to find the latent shared features for B-mode and RF time series data in $D^S_L$. We train a Support Vector Machine (SVM) classifier with Radial Basis Function (RBF) kernel. SVM has been previously demonstrated to differentiate between various tissue types using TeUS data in *ex vivo* [107] and *in vivo* [13, 76] studies with high accuracy. In addition to the binary output of the classifier, we estimate the likelihood of the ROIs to be cancerous [13]. More detail about SVM can be found at Section 3.2.

5.2.2.3 Baseline Classification

In addition to domain adaptation as a solution for the accessibility of *in vivo* TeUS RF data for prostate cancer detection, we explore other plausible solutions as baselines for comparative analysis versus the proposed method. These include direct deployment of the trained DBN network with TeUS RF data in [11, 13] and test on TeUS B-mode data (BL-1), and train a DBN model using the method proposed in [11] with TeUS B-mode data and test on the same data type (BL-2).

5.2.2.4 Generalization

To remove any bias that could be potentially introduced in training/adaptation, we randomly divide the data into independent subsets of training, validation, and testing. Specifically, the test data was generated by fully excluding data of half of the biopsy cores from $D^S_L$ and $D^S_U$. This new introduced data division include, 42 biopsy cores in the test dataset, $D^S_{test}$. For the remaining 42 cores, we will follow the procedure that we explain earlier for training and validation.

5.2.3 Results and Discussion

To evaluate the proposed approach, we conducted a series of experiments to assess the performances of the unsupervised domain adaptation and classification.
5.2. Transfer Learning From TeUS RF to B-mode Spectral Features

5.2.3.1 Unsupervised Domain Adaptation

**DBN Parameters:** As we mentioned earlier, to find a common feature space between RF and B-mode time series, we train a shared DBN between $\mathcal{D}_{US}^S$ and $\mathcal{D}_{UT}^S$. We used the MATLAB library of Tanaka et al. [146] to build and train the DBN. The training step requires setting the values of numerical meta-parameters such as the learning rate (LR), momentum, weight-cost, initial values of the weights, number of hidden units, and mini-batch size (BS). For this purpose, we randomly divide $\mathcal{D}_U^S$ to training (80%) and validation (20%) datasets. Initially, we set the DBN structure including the number of hidden layers and nodes configuration, as well as the numerical meta-parameters to the default values of the DBN library [146]. The DBN structure consists of two RBM layers with $d = 11$ visible and hidden units in the first and last layers. In the proposed architecture, we need to set the number of hidden neurons for the first layer, $n$ (see Fig. 5.1), as well as the learning rate and mini-batch size. The purpose of training is to reduce error to a reasonably low value in as few epochs as possible. We heuristically search for $n$, so that cross entropy between a sample and its reconstruction through DBN is minimized in both the training and validation datasets in as few epochs as possible. Figure 5.2(a) shows the cross-entropy loss function for different numbers of hidden neurons using 250 iterations in pre-training and training steps. Since the lowest error is obtained with $n = 44$, the finalized DBN is composed of a real-valued visible layer with 11 units, and two hidden layers consisting of 44 and 11 hidden units, respectively. The learning rate and mini-batch size are set similarly with a coarse search as shown in Fig. 5.2. Figure 5.2(b) shows the cross-entropy loss function for different numbers of hidden neurons using 250 epochs, where we have the lowest errors for the $LR = 0.2$. As it is shown in Fig. 5.2(c), we also achieved the lowest error for the $BS = 10$. The momentum and the weight cost values do not change from the default values ($0.9, 2 \times 10^{-4}$). In the fine-tuning step, we ran 250 epochs with a fixed learning rate of 0.2, and a mini-batch size of 10. After completion of the learning procedure, the last hidden layer of the DBN produces the latent common feature representation.

**Domain Alignment:** Figure 5.3 shows the divergence between RF and B-mode time series data in $\mathcal{D}_L^S$ prior to (top row) and after (bottom row) domain alignment using the shared DBN. This figure depicts a normal distribution which fitted to the histogram of the top three features obtained following PCA whitening. It is obvious that the proposed domain adaptation method can effectively align features in common learned feature space. Moreover,
5.2. Transfer Learning From TeUS RF to B-mode Spectral Features

![Learning Curve for DBN Training](image)

Figure 5.2: Learning curve for DBN training based on the cross-entropy: (a) for first hidden layer size. (b) for different learning rates (LR). (c) for different mini-batch size (BS). In a coarse search for the meta-parameters we achieved the lowest cross entropy loss with $n = 44$, $LR = 0.2$, and $BS = 10$.

we use the Subspace Disagreement Measure (SDM) \cite{54} to evaluate the domain difference \cite{30, 46}. In a subspace with dimension $d$, the domain difference is calculated as $\Delta D = \sum_{i=1}^{d} SDM(i)$. For perfectly aligned feature distributions, $SDM(i) = 0$. So, the lower the $\Delta D$ is, the better the domains are aligned. $\Delta D$ decreases from 0.708 to 0.449 for RF and B-mode time series data in $d = 11$ feature space as a result of domain adaptation.

Aligning domains helps compensate for the domain distribution shift, thus making the decision models trained on the aligned source space comparable with the source-aligned target domain. As a result, we can validate the performance of our domain adaptation method indirectly using the classification accuracy obtained in supervised training step with aligned source and target data.

5.2.3.2 Supervised Classification

**Data Division:** To determine the sensitivity of our methodology to the choice of the training labeled-data, we create different partitions of training and testing datasets from $D_s^L$ in a hold-out validation setting. For each new pair of datasets, we randomly select $x\%$ of cores in the labeled dataset as the training data for the supervised training of the SVM and the other $100\% - x\%$ cores are used as the testing data. We generate three datasets including RF and B-mode time series data with $x = 50\%, 75\%$, and $85\%$ of the total 84 biopsy cores in $D_s^L$.

**Classification Performance:** We assess the overall performance of our approach using the AUC. This curve depicts relative trade-offs between
5.2. Transfer Learning From TeUS RF to B-mode Spectral Features

Figure 5.3: Distribution shift from B-mode to RF for the top three features before (top row) and after (bottom row) the shared deep network. The proposed domain adaptation method can effectively align features and reduce the distribution shift in common learned feature space.

sensitivity and specificity. The larger AUC values indicate better classification performance. Accuracy, sensitivity, and specificity are also measured for the test data. For this purpose, we consider all of the cancerous cores as the positive class and all the benign cores as the negative class. Sensitivity is the percentage of cancerous cores that are correctly identified as cancerous compared to the pathology results; specificity is the proportion of non-cancerous cores that are correctly classified, and accuracy is the percentage of true results (both true positives and true negatives) in the total number of cores. For each data division, we randomly select 10 different training and testing datasets and report the average performance. Table 5.1 and Table 5.2 show the classification results for differentiating between cancerous and benign cores.

Size of the Labeled Dataset: We investigate how the learning curve progresses with increasing the number of training samples. This is indirectly assessed with $k$-fold cross-validation increasing $k$ from 2 to 6. Figure 5.4 and Table 5.1 show AUC, accuracy, sensitivity and specificity of the classification
5.2. Transfer Learning From TeUS RF to B-mode Spectral Features

Table 5.1: Model performance measured by AUC for classification in different data divisions.

<table>
<thead>
<tr>
<th>Data division</th>
<th>RF time series (±)</th>
<th>B-mode series (±)</th>
</tr>
</thead>
<tbody>
<tr>
<td>50% train, 50% test</td>
<td>0.77 (± 0.01)</td>
<td>0.68 (± 0.01)</td>
</tr>
<tr>
<td>75% train, 25% test</td>
<td>0.72 (± 0.03)</td>
<td>0.70 (± 0.03)</td>
</tr>
<tr>
<td>85% train, 15% test</td>
<td>0.71 (± 0.01)</td>
<td>0.71 (± 0.02)</td>
</tr>
</tbody>
</table>

Table 5.2: Model performance measured by specificity and sensitivity for classification in different data divisions.

<table>
<thead>
<tr>
<th>Data division</th>
<th>RF time series</th>
<th>B-mode series</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Specificity</td>
<td>Sensitivity</td>
</tr>
<tr>
<td>50% train, 50% test</td>
<td>0.69 (± 0.03)</td>
<td>0.70 (± 0.02)</td>
</tr>
<tr>
<td>75% train, 25% test</td>
<td>0.69 (± 0.01)</td>
<td>0.79 (± 0.07)</td>
</tr>
<tr>
<td>85% train, 15% test</td>
<td>0.75 (± 0.06)</td>
<td>0.78 (± 0.06)</td>
</tr>
</tbody>
</table>

for different cross-validation settings. A two-way t-test between $k = 2$ and $k = 6$ fails to show statistically significant difference in AUC, accuracy, sensitivity and specificity ($p > 0.05$).

**Integration of TeUS and mp-MRI:** In another analysis, we combine the TeUS cancer detection results with readings from mp-MRI. The combination method takes advantage of both imaging modalities. If mp-MRI declares cancer suspicious level as low or high (see Section 2.1) for a core, we use its predictions alone and declare the core as benign or aggressive cancer, respectively. On the other hand, when mp-MRI declares the cancer suspicious level as intermediate, we use predictions based on TeUS data. We evaluate the combined approach with the data division explained earlier in Section 5.2.3.2. Table 5.3 and Table 5.4 show the classification performance for differentiating between cancerous and benign cores using the combined method measured by AUC, specificity, and sensitivity. The combined approach leads to an AUC of 0.81 for predicting cancer versus benign cores using RF time series and AUC of 0.79 using B-mode time series.

5.2.3.3 Baseline Classification

We assess the comparative performance of the proposed method over the baselines (Section 5.2.2.3) in terms of AUC, sensitivity, and specificity. The performance evaluating the metrics are discussed earlier in the section.
5.2. **Transfer Learning From TeUS RF to B-mode Spectral Features**

![Graph](image)

**Figure 5.4:** Influence of labeled dataset size in classification accuracy: performance of the method measured by AUC, accuracy, sensitivity and specificity in the k-fold cross-validation setting for (a) TeUS RF data and (b) TeUS B-mode data.

**Table 5.3:** Performance for the combination of mp-MRI and TeUS measured by AUC for classification in different data divisions.

<table>
<thead>
<tr>
<th>Data division</th>
<th>RF time series</th>
<th>B-mode series</th>
</tr>
</thead>
<tbody>
<tr>
<td>50% train, 50% test</td>
<td>0.80 (± 0.01)</td>
<td>0.76 (± 0.01)</td>
</tr>
<tr>
<td>75% train, 25% test</td>
<td>0.81 (± 0.07)</td>
<td>0.77 (± 0.03)</td>
</tr>
<tr>
<td>85% train, 15% test</td>
<td>0.81 (± 0.06)</td>
<td>0.79 (± 0.01)</td>
</tr>
</tbody>
</table>

Table 5.5 shows the model performance in different data divisions using **BL-1**, **BL-2**, TeUS RF and adapted TeUS B-mode approaches. As expected, for **BL-1** where we directly use B-mode data with a model trained using RF data, the AUC is close to a random prediction. These observations show that, while TeUS RF and B-mode data are related, the differences in their corresponding feature distributions are significant and affect classification results. Moreover, for **BL-2** where we used TeUS B-mode data for training and testing of the DBN model, the obtained AUC is 0.58. The results show that the obtained AUCs are significantly below that of TeUS RF and TeUS B-mode data (p<0.05 for either **BL-1** and **BL-2** versus TeUS RF.).

**5.2.3.4 Generalization**

For the 42 cores that we explained in Section 5.2.2.4, we follow the procedure mentioned earlier for training and validation. Table 5.6 shows performance of the proposed approach in $D_{test}^S$. On this test set, the classification resulted in an AUC of 0.71 for predicting prostate cancer using only TeUS RF data and
5.2. Transfer Learning From TeUS RF to B-mode Spectral Features

Table 5.4: Performance for the combination of mp-MRI and TeUS measured by specificity and sensitivity for classification in different data divisions.

<table>
<thead>
<tr>
<th>Data division</th>
<th>RF time series</th>
<th>B-mode series</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Specificity</td>
<td>Sensitivity</td>
</tr>
<tr>
<td>50% train, 50% test</td>
<td>0.73 (± 0.03)</td>
<td>0.75 (± 0.01)</td>
</tr>
<tr>
<td>75% train, 25% test</td>
<td>0.77 (± 0.07)</td>
<td>0.83 (± 0.04)</td>
</tr>
<tr>
<td>85% train, 15% test</td>
<td>0.80 (± 0.07)</td>
<td>0.78 (± 0.04)</td>
</tr>
</tbody>
</table>

Table 5.5: Comparison of model performance measured by AUC using baselines and the proposed approach in different data divisions.

<table>
<thead>
<tr>
<th>Data division</th>
<th>BL-1</th>
<th>BL-2</th>
<th>RF series</th>
<th>B-mode series</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>AUC</td>
<td>Specificity</td>
<td>AUC</td>
<td>Specificity</td>
</tr>
<tr>
<td>50% train, 50% test</td>
<td>0.50 (± 0.08)</td>
<td>0.58 (± 0.06)</td>
<td>0.77 (± 0.01)</td>
<td>0.69 (± 0.02)</td>
</tr>
<tr>
<td>75% train, 25% test</td>
<td>0.49 (± 0.10)</td>
<td>0.61 (± 0.03)</td>
<td>0.72 (± 0.02)</td>
<td>0.70 (± 0.01)</td>
</tr>
<tr>
<td>85% train, 15% test</td>
<td>0.51 (± 0.12)</td>
<td>0.60 (± 0.02)</td>
<td>0.71 (± 0.01)</td>
<td>0.71 (± 0.03)</td>
</tr>
</tbody>
</table>

0.70 using domain adaptation with TeUS B-mode data. When integrating TeUS with MRI readings, these results were increased to AUCs of 0.76 and 0.74, respectively.

Figure 5.5 shows the comparative performance, measured by AUC, of the proposed method over the baselines (Section 5.2.2.3) for \( \mathcal{D}^{S}_{test} \). For BL-1, the AUC = 0.51 and it is close to a random prediction. In comparison to the TeUS RF approach, we can see a significant improvement \( (p < 0.05) \). For BL-2, AUC = 0.59, which is significantly below that of either TeUS RF or TeUS B-mode data \( (p < 0.05) \).

Table 5.6: Performance for the TeUS only and combination of mp-MRI and TeUS measured by AUC, specificity, and sensitivity for classification in the test dataset.

<table>
<thead>
<tr>
<th>Data division</th>
<th>RF time series</th>
<th></th>
<th>B-mode series</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>AUC</td>
<td>Specificity</td>
<td>Sensitivity</td>
<td>AUC</td>
</tr>
<tr>
<td>TeUS</td>
<td>0.71</td>
<td>0.73</td>
<td>0.71</td>
<td>0.70</td>
</tr>
<tr>
<td>TeUS + mp-MRI</td>
<td>0.76</td>
<td>0.78</td>
<td>0.76</td>
<td>0.74</td>
</tr>
</tbody>
</table>
5.3. Transfer Learning From TeUS RF to B-mode Using RNN

Figure 5.5: The comparative performance of the proposed method measured by AUC over the baselines for $D_{test}^S$

5.2.3.5 Colormaps:

Figure 5.6 shows examples of the cancer likelihood maps, derived from the output of SVM using both TeUS RF and B-mode data, overlaid on B-mode US image. We use the approach described in our earlier publication [11] for this purpose. In the colormaps, red regions belong to ROIs for which the cancer likelihood is more than or equal to 50%.

5.3 Transfer Learning From TeUS RF to B-mode Using RNN

Towards a real-time system for temporal enhanced ultrasound-guided prostate biopsy, in this section, we propose an RNN-based transfer learning solution which is solely working with TeUS B-mode data streaming from an ultrasound scanner. The method is implemented as a part of a unified software framework demonstrating near real-time analysis of ultrasound data stream using a deep learning solution. The detail description and additional information about the software framework are presented in Appendix B.

At the center of the proposed software solution, a probabilistic model is responsible for analysis of temporal enhanced ultrasound using RNN. The model is trained to capture tissue-dependent feature from TeUS B-mode data which rely on the successful knowledge transfer between TeUS RF and TeUS B-mode data in the unsupervised training step. The transfer learning and diagnosis approach is evaluated for cancer detection accuracy on ultrasound data obtained from a large clinical study with 255 biopsy cores from 157
5.3. Transfer Learning From TeUS RF to B-mode Using RNN

![Figure 5.6: Cancer probability maps overlaid on B-mode US image, along with the projected needle path in the temporal US data and centered on the target. The ROIs for which the cancer likelihood is more than 50% are colored in red, otherwise they are colored as blue. The red boundary shows the segmented prostate in MRI projected in TRUS coordinates, dashed line shows needle path and the arrow pointer shows the target: (a)-(c) Correctly identified cancerous core using RF time series data; (b)-(d) Correctly identified cancerous core using B-mode time series data.](image)

subjects. The proposed solution is further assessed with an independent dataset with 21 biopsy targets from six subjects. In the first study, we achieve the area under the curve, sensitivity, specificity, and accuracy of 0.94, 0.77, 0.94 and 0.92, respectively, for the detection of prostate cancer. In the second study, we achieve an AUC of 0.85. Our results suggest that TeUS-guided biopsy can be potentially effective for the detection of prostate cancer.

5.3.1 Materials

The deep networks are generated using a dataset consisting of biopsy targets in mp-MRI-TRUS fusion-biopsies with 255 biopsy cores from 157 subjects. We refer to this data as the first retrospective study. As we explained before, for development of our transfer learning system, we use both unlabeled data.
from the whole ultrasound scan and labeled data from the biopsy target region coming from the dataset explained in Chapter 2. In this section, we also performed another small independent prostate biopsy study for further assessment of our approach. The details of materials and the complementary study can be found in following sections.

5.3.1.1 Data Division

The data from 255 biopsy cores is divided into mutually exclusive training, $D_{train}^T$, and test sets, $D_{test}^T$. Training data, $D_{train}^T$, is made up of 84 cores from patients with homogeneous tissue regions which we randomly select them. We further use the test data, $D_{test}^T$, consists of 171 cores to evaluate the trained model during the guidance system implementation, where 130 cores are labeled as benign and 31 cores are labeled as cancerous with GS $\geq 3+3$.

Unlabeled Data: As we mentioned in Section 2.3.1, in temporal representation of TeUS, each ROI is represented by $T = 100$ echo intensity value over 100 consecutive ultrasound frames as $x = (x_1, ..., x_T)$. Unlike our typical biopsy region analysis, here, the RF time series data from the ROIs of the entire imaging plane of the prostate and in the scan level make up the unlabeled source dataset, $D_{US}^T$, while the corresponding TeUS B-mode sequence make up the unlabeled target dataset, $D_{UT}^T$, where $D_U^T = D_{UT}^T \cup D_{US}^T$. Both $D_{US}^T$ and $D_{UT}^T$ have an equal number of samples. In total, $D_U^T$ includes 5,255,680 unlabeled training samples, consisting of 2,627,840 RF and 2,627,840 B-mode time series data. To maintain the mutual exclusiveness between the train and test data, for generation of above unlabeled data, we only use the cores in $D_{train}^T$.

Labeled Data The labeled source dataset, $D_{LS}^T$, includes TeUS RF sequence of ROIs and the target dataset $D_{LT}^T$, includes the corresponding TeUS B-mode sequence. Given the data augmentation strategy, we obtain a total number of 129,024 training samples for both TeUS RF and TeUS B-mode data of the cores that belong to $D_{train}^T$.

5.3.1.2 Complementary Second Retrospective Study

To further assess the developed solution, we use the data that we acquired in the second retrospective study. As we explained in Section 2.4 we performed a second independent fusion biopsy study including six subjects. In this study, only TeUS B-mode data were recorded for each target to minimize disruption
5.3. Transfer Learning From TeUS RF to B-mode Using RNN

Table 5.7: Gleason score distribution in the second retrospective clinical study.

<table>
<thead>
<tr>
<th>GS</th>
<th>Benign</th>
<th>GS 3+3</th>
<th>GS 3+4</th>
<th>GS 4+3</th>
<th>GS 4+4</th>
<th>GS 4+5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of Cores</td>
<td>7</td>
<td>2</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>6</td>
</tr>
</tbody>
</table>

to the clinical workflow. This study resulted in 21 targeted biopsy cores with GS distribution as explained in Table 5.7. We use the histopathology labeling of the cores as the ground-truth to assess the accuracy of the guidance system in detecting the cancerous lesions.

5.3.2 Methods

The deep networks are generated mainly based on the methods that we presented in our earlier works and previous sections. We give a brief overview of these methods. For a detailed description of the models, the reader may refer to Section 3.2 and Section 5.2 [10, 14].

An individual TeUS sequence of length \( T \), \( x^{(i)} \), is composed of echo-intensity values \( x_t^{(i)} \) for each time step, \( t \), and is labeled as \( y_t^{(i)} \in \{0,1\} \), where zero and one indicate benign and cancer biopsy outcome, respectively. We aim to learn a mapping from \( x_t^{(i)} \) to \( y_t^{(i)} \) in using RNNs to model the temporal information in TeUS B-mode data.

For this purpose, first, we use the plain version of the unsupervised domain adaption approach presented in [14] to find a common feature space between TeUS RF and TeUS B-mode data. Let \( x_t^{(i)} \) indicates the \( i^{th} \) RF TeUS sample in \( D_{TS} \) and \( x_t^{(i)} \) is the corresponding \( i^{th} \) B-mode TeUS sample in \( D_{UT} \). Also, \( h_t^{(i)} \) indicate the output of the RNN for \( x_t^{(i)} \). Here, we use a layer of RNN with Long Short-Term Memory (LSTM) cells [50], thus, the size of input and output is equal to \( T = 100 \). Based on the Equation 5.1 in Section 5.2 to find the common feature space between the source and domain, the unsupervised training stage minimizes the cross-entropy errors as the loss function, \( L_{Div} \):

\[
L_{Div} = - \sum_{i \in D_{UT}} x_t^{(i)} \log h_t^{(i)},
\]

Given an input TeUS B-mode sequence \( x_t^{(i)} = (x_1^{(i)}, \ldots, x_T^{(i)}) \), our RNN computes a hidden vector sequence \( h_t^{(i)} = (h_1^{(i)}, \ldots, h_T^{(i)}) \) in the common
5.3. Transfer Learning From TeUS RF to B-mode Using RNN

feature space between the RF and B-mode data within a domain adaption step. After the unsupervised domain adaptation step, our discriminative RNN architecture is built with LSTM cells [50], where each cell, maintains a memory over time. We use two layers of LSTMs with $T = 100$ hidden units to capture temporal changes in data. Following these layers, we use a fully connected layer to map the learned sequence to a posterior over binary classes of benign and cancer tissue in a supervised classification step. This final node generates a predicted label, $\mathbf{y}^{(i)}$, for a given TeUS B-mode ROI sequence, $\mathbf{x}^{(i)}_b \in \mathcal{D}_{LT}^T$. The training criterion for the network is to minimize the loss function as the binary cross-entropy between $y^{(i)}$ and $\mathbf{y}^{(i)}$ over all of the training samples where we use Root Mean Square Propagation (RMSprop) optimizer. As recommended, using a training-validation setting, we perform a grid search to find the optimum hyperparameter in our search space. Once the optimum hyperparameters are identified (the number of RNN hidden layers = 2, batch size = 64, learning rate = 0.0001, dropout rate = 0.4, and the regularization term = 0.0001), the entire training set, $\mathcal{D}_{train}^T$, is used to learn the final classification model.

5.3.3 Results and Discussion

5.3.3.1 Classification model validation

To evaluate the proposed RNN-based approach, we used clinical data from two retrospective studies to simulate the flow of information across different modules. To validate the accuracy of the classification model, we use $\mathcal{D}_{test}^T$. We use sensitivity, specificity, and accuracy in detecting cancerous tissue samples to report the validation results. We consider all cancerous cores as the positive class and other non-cancerous cores as the negative class. Sensitivity is defined as the ratio of cancerous cores that are correctly identified while specificity is the ratio of non-cancerous cores that are correctly classified. Accuracy is the ratio of the correctly identified results over the total number of cores. We also report the overall performance of our approach using the AUC. Table 5.8 shows the model performance for classification of cores in $\mathcal{D}_{test}^T$ for different MR suspicious levels. For samples of moderate MR suspicious level (70% of all cores), we achieve an AUC of 0.94 using the LSTM-RNN. In this group, our sensitivity, specificity, and accuracy are 0.75, 0.96, and 0.93, respectively. In comparison, only 26% of all of the cores identified in mp-MRI are cancerous after biopsy which means our approach can effectively complement mp-MRI during the guidance procedure to reduce the number of false positives for those targets with moderate MR suspicious
5.3. Transfer Learning From TeUS RF to B-mode Using RNN

Figure 5.7: Guidance interface implemented as part of a 3D Slicer module: cancer likelihood map is overlaid on B-mode ultrasound images. Red indicates predicted labels as cancer, and blue indicates predicted benign regions. The boundary of the segmented prostate is shown with white and the green circle is centered around the target location which is shown in the green dot.

Table 5.8: Model performance for classification of cores in the test data from the first retrospective study for different MR suspicious levels. \( N \) indicates the number of cores in each group.

<table>
<thead>
<tr>
<th>Method</th>
<th>Specificity</th>
<th>Sensitivity</th>
<th>Accuracy</th>
<th>AUC</th>
</tr>
</thead>
<tbody>
<tr>
<td>All of the biopsy cores ( (N = 171) )</td>
<td>0.94</td>
<td>0.77</td>
<td>0.92</td>
<td>0.94</td>
</tr>
<tr>
<td>Moderate MR suspicious level ( (N = 115) )</td>
<td>0.96</td>
<td>0.75</td>
<td>0.93</td>
<td>0.94</td>
</tr>
<tr>
<td>High MR suspicious level ( (N = 20) )</td>
<td>0.85</td>
<td>0.98</td>
<td>0.95</td>
<td>0.95</td>
</tr>
</tbody>
</table>

5.3.3.2 System assessment

Figure 5.7 shows the guidance interface implemented as part of a 3D Slicer module running on the client machine. In order to evaluate the performance of the proposed solution, other than the subjective evaluation of guidance visualization, the accuracy of the target detection and the run-time are measured.

Guidance accuracy: To further assess the developed approach, we performed a second independent MRI-TRUS fusion biopsy study. This study
Table 5.9: Run-time of the steps of the prostate guidance system averaged over N = 21 trials with data from the second retrospective study (given as mean±std).

<table>
<thead>
<tr>
<th>Host Machine</th>
<th>Operation</th>
<th>Average time (s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>TeUS-Client</td>
<td>Guidance Visualization</td>
<td>0.40 (± 0.11)</td>
</tr>
<tr>
<td>TeUS-Server</td>
<td>Classification</td>
<td>1.66 (± 0.32)</td>
</tr>
<tr>
<td></td>
<td>Segmentation</td>
<td>0.12 (± 0.17)</td>
</tr>
</tbody>
</table>

resulted in 21 targeted biopsy cores with GS distribution as explained in Table 5.7. We achieve an AUC, sensitivity, specificity, and accuracy of 0.85, 0.93, 0.72 and 0.85, respectively, for the fusion biopsy targets. Our results show that the only miss-classified cancerous target is GS 3+3 with the tumor in core less than 0.4 cm.

Run-time: The run-time was measured for all parts of the workflow (classification, segmentation, and guidance visualization) using a timer log provided by the open-source Visualization Toolkit. The TeUS-client computer featured an 2.70 GHz Intel Core™ i7-6820 (8 CPUs) processor with 64 GB of RAM, Windows 10 Enterprise N, Visual Studio 2015, 3D Slicer 4.3.1, OpenMP 2.0, and Intel MKL 11.0. The TeUS-server was hosted by a computer running an Ubuntu 16.04 operating system and a 3.4 GHz Intel Core™ i7 CPU with 16 GB of RAM, equipped with GeForce GTX 980 Ti GPU with 2816 CUDA cores and 6 GB of memory.

Run-time results are summarized in Table 5.9. Averaged over all N = 21 trials in the second fusion biopsy study, the total time spent on TeUS-client for guidance visualization and post-processing is 0.40 ± 0.11 seconds. Since the classification and segmentation modules run simultaneously, the TeUS-server run-time is constrained by the classification task run-time of a batch of 100 ultrasound data frames, which is 1.66 ± 0.32 second. Within the context of prostate biopsy guidance workflow, this means an addition of only about 20 sec to a 20 min procedure. We consider this near real-time performance sufficient for the requirements of this clinical procedure.

5.3.4 Discussion and comparison with other methods
The best result to-date involving TeUS B-mode data (AUC = 0.7) is based on spectral analysis using DBN as the underlying machine learning framework [11, 13]. Comparing our LSTM-RNN approach with that method as
5.4 Conclusion

In this chapter, we tried to address the challenge of accessibility of RF data in the commercial ultrasound scanners. In the first part, we proposed a novel method for unsupervised domain adaptation of RF (source) and B-mode (target) time series data. We worked towards the objective of making the method maximally discriminative for prostate cancer detection on both source and target domain by leveraging: (1) a shared deep network trained using relatively large amount of unlabeled TeUS data from source and target, and (2) a shared classifier using a few labeled examples from both domains. We demonstrated that in presence of a distribution shift between RF and B-mode data, a transfer learning method can compensate for the divergence between the distribution and enable TeUS B-mode-based tissue classification as an alternative to TeUS RF-based approach for the prostate cancer detection, where RF data may not be easily accessible on a
commercial ultrasound scanner. The need for domain adaptation is supported by a recent paper [30], which also reported that even within RF data mode, small changes to the scattering environment result in significant shifts in the distribution of features. The proposed approach is successful in factoring out the effect of domain shift in TeUS data by adapting cross-entropy loss between the two domains at the common shared feature space. In an in vivo study including 84 biopsy targets, we achieved the AUC of 0.70 and accuracy of 0.73 using TeUS B-mode data. By using TeUS RF data, the AUC and accuracy are 0.72 and 0.73, respectively.

We considered this study, as a feasibility investigation to take advantage of knowledge transfer methods in ultrasound-based diagnosis and intervention techniques. This approach can be used as a versatile pre-trained network to reduce divergence between these two domains, which can be fine-tuned for a given clinical task. An extension of the approach has been used in the second part, to capture variations in TeUS RF and B-mode data in real-time using temporal analysis of TeUS with RNNs. We have to demonstrate that accuracy of the integrated system on clinical ultrasound devices that only provide B-mode data is comparable to those in bench-top prototypes in the laboratory. The system was validated using retrospective in vivo datasets including both TeUS RF and B-mode data of 255 biopsy target cores obtained in a large clinical study during mp-MRI-TRUS fusion-biopsy. For the validation data, we achieved an AUC of 0.94, and sensitivity and specificity of 0.77, and 0.94 respectively. The integrated system was then evaluated with an independent in vivo dataset obtained during mp-MRI-TRUS fusion-biopsy from six subjects and includes only B-mode TeUS of 21 biopsy targets. We achieved an AUC of 0.85 for this dataset, and sensitivity and specificity of 0.93 and 0.72, respectively. The average sensitivity of the system considering both studies is 82% where 18% of cancerous cores (10 out of 55) were not identified.

From a clinical perspective, the motivation of the work is to enable real-time assessment of TeUS for prostate cancer detection. The immediate goal of the current work is to demonstrate the viability of this approach using retrospective data with known ground-truth so that we can optimize the system’s performance and understand how such system would fit within the standard clinical workflow. The promising results of this initial assessment indicate that our proposed TeUS-based system is capable of providing guidance information for the prostate biopsy procedure. Future work should focus on prospective evaluation and feasibility assessment of the biopsy guidance system.
Chapter 6

Investigation of Physical Phenomena Underlying TeUS

The particular aspect of time that I’m interested in is the arrow of time: the fact that the past is different from the future. We remember the past but we don’t remember the future.

— Sean M. Carroll

6.1 Introduction

Temporal Enhanced Ultrasound (TeUS) is a novel non-invasive imaging paradigm that captures information from a temporal sequence of backscattered ultrasound Radio-Frequency (RF) or B-mode image data obtained from a fixed tissue location. This technology has been shown to be effective for classification of various in vivo \[11, 13, 73, 75, 109\] and ex vivo \[107, 108\] tissue types including prostate cancer from benign tissue. A comparison of TeUS with the analysis of power spectrum of RF data for tissue characterization showed that TeUS and RF spectral analysis compliment each other \[73, 107\],

with TeUS showing higher overall AUC. More recently, as we have shown in previous chapters, TeUS RF and B-mode can successfully be used for diagnosis and characterization of prostate cancer and its extent [9–14, 16].

While the physical phenomenon governing temporal ultrasound/tissue interaction is the subject of ongoing investigation in our group, several hypotheses have been explored so far. Our previous studies have indicated two primary phenomena that influence TeUS:

1. Changes in tissue temperature due to acoustic absorption: It has been proposed that the acoustic radiation force of the transmit US signal increases the temperature and changes the speed of sound in different tissue types [36].

2. Micro vibrations of tissue due to physiological vibration: It has also been suggested that a combination of micro-vibration of acoustic scatters in micro-structures and the density of cells play a role [110].

To optimize TeUS for clinical translation, we aim to investigate the physical processes that govern US-tissue interaction during the acquisition of TeUS data [108]. We previously investigated changes in tissue temperature during TeUS data acquisition, using a numerical model [36]. The results demonstrated that changes in tissue temperature, which affect the speed of sound, can be used for tissue characterization. However, even with exaggerated image settings of high frame rate and acoustic power that do not match clinical imaging conditions, classification results were substantially below those achieved in our ex vivo and in vivo studies [36].

Our results from Chapter 3 and Chapter 4 showed a consistently high classification accuracy in a large dataset in this thesis. The preliminary feature visualization results from Chapter 3 shows that frequencies between 0 – 2 Hz provide the most discriminative features for distinguishing cancerous and benign tissue. These results suggest that the phenomenon is consistent for the two independent training and test datasets in clinical settings. Interestingly, the range of frequencies that we have identified as most discriminative between cancerous and benign tissue (0 – 2 Hz in Fig. 6.5) are also consistent with the ranges we have observed in our previous independent studies [75, 76].

In this section, we investigate the second phenomenon related to tissue micro vibrations as the main source of the tissue typing capabilities of TeUS with clinical image settings. Possible sources of tissue micro vibration include external, low-amplitude environmental vibrations, and internal physiological motion such as pulsation due to the heart beat [10, 19]. In design and
6.2 Spectral Feature Visualization

In this section, we use the data from the first study as we explained in Chapter 2 and we focus on the spectral representation of TeUS. From the whole 255 cores in our dataset, we divide the data from 197 cores, into the train, \( D_{train}^S \), and test, \( D_{test}^S \), sets as explained in Section 4.2. For building the trained DBN model, we use \( D_{train}^S \) along with the distribution learning approach that we explained in Section 4.2. The test data, \( D_{test}^S \), consists of 165 cores from 114 patients, with the following distribution: 121 benign, 12 GS of 3+3, 14 GS of 3+4, 2 GS of 4+3, and 16 GS of 4+4.

6.2.2 Methodology

To determine the characteristics of the non-cancerous and cancerous tissue samples in the TeUS data and their correlation with the learned features, we propose a feature visualization approach for the network that we discussed in Section 4.2. We use this approach to identify the most discriminative frequency components of the TeUS RF data as learned in the feature learning step (See Sections 4.2.2.1). Figure 6.1 shows an overview of the proposed method. First, TeUS test data, \( D_{test}^S \), is propagated through the trained DBN, and the activations of the last hidden layer (i.e., the learned latent features) are computed. We then use the GMM along with the learned features, as explained in Section 4.2 to assign a prostate cancer grades to each ROIs of the test dataset. To examine the significance of the \( i^{th} \) \((i = 1, ..., 6)\) individual learned latent feature in the detection of different grades, the activations of all other five hidden units in the third layer are set to zero. The activation of the non-zero learned feature is back-propagated to the input layer. The
resulting reconstructed signal, displayed in the input layer as a series of frequency components, highlights those components that contribute to the activation of the non-zero learned feature. By comparing the components activated for ROIs labeled as GS pattern of 3, 4 as well as non-cancerous tissue in Section 4.2, we can identify those frequency ranges that are different between two tissue types. This process is performed for all the six latent features.

The results of feature visualization (Fig. 6.5(a)) suggest that frequencies between $0 - 2 \, Hz$ provide the most discriminative features for distinguishing between Gleason pattern 3 and 4 as well as non-cancerous tissue samples (See Section 6.4). Moreover, the identified frequency range is also consistent with the ranges that we have observed in our previous independent studies [11, 75, 76]. Our results to-date suggest that tissue micro-vibration, possibly due to pulsation from the heartbeat ($\sim 1.2 \, Hz$), is a key contributor to the tissue typing capability of TeUS. To further examine this hypothesis, we continue our study with a histopathology mimicking simulation as explained in the next section.
6.3. Histopathology Mimicking Simulation

A primary source of the observed backscattered ultrasound signal from the tissue has been associated with the scattering from the cell nuclei [8, 69]. Prostate cancer primarily presents as changes in tissue microstructure where different density, size and spatial arrangement of nuclei are observed [35, 71]. Subtle differences in scattering distribution results in significant changes in the back-scattered signal [68]. We hypothesize that the induced tissue micro-motion due to blood pulsation can cause different US scattering patterns in various microstructures of the prostate, which in turn can be captured by TeUS for prostate cancer grading. Figure 6.2 shows an overview of our simulation design to explore this effect.

6.3.1 Digital Pathology Data

We use a digital pathology dataset [71] to investigate this hypothesis in a histopathology-based simulation framework. Our dataset includes 14 digital histology slides of prostate cancer patients [71]. We extract the positions of nuclei from these slides where the locations of cancerous cells are marked by an expert pathologist. We divide the digitized slides into blocks of $2 \times 2 \, mm \times mm$, with a resolution of $0.5 \, \mu m/pixel$. Here, a single scattering point is considered at each nucleus location segmented from the pathology slides.

6.3.2 Numerical Simulation Design

For the simulation of mechanical micro-vibrations, we design tissue-mimicking phantoms which are generated using digital histology slides. We extract the locations of nuclei in the whole-mount digital slides. As we mentioned before, we divide each slide into blocks of $2 \times 2 \, mm \times mm$. We then generate a 2D
6.3. Histopathology Mimicking Simulation

Figure 6.3: ROI selection and nuclei-based scatterer generation process: (a) Sample of histopathology slide [70], where the red boundary depicts the cancer area; (b) digitized slide overlaid on the histopathology slide, where green and red areas represent the benign and cancer regions, respectively. The selected ROIs are shown by black squares; (c) extracted nuclei positions in the selected ROIs; left: a cancer region, right: a benign region; (d) the extracted positions of nuclei from each ROI is embedded in an FEM model.

model of nuclei positions based on their coordinates in the blocks. A total number of 42 blocks including 14 with GS≥4, 14 with GS=3 and 14 from non-cancerous tissue types is selected. We place each block at the center of a homogeneous linear viscoelastic phantom as shown in Fig. 6.2 and Fig. 6.3. The size of the phantom is $5 \times 5 \times 5 \text{ cm}^3$, and the elasticity and viscosity are set to 25 kPa and 2.15 Pa.s, respectively [92].

To capture micro-motion due to an external excitation source, we use a Finite Element Model (FEM) simulation in COMSOL Multiphysics 5.2 (Burlington, MA, USA). The external vibration source is defined as a sinusoidal wave with the frequency of 1 Hz and amplitude of 100 µm. We pulsate the inferior surface of the defined tissue-mimicking phantom with this low-frequency signal. The cells’ nuclei, which are embedded in the tissue phantom, are displaced in our FEM as a result of the applied mechanical vibration.

6.3.3 TeUS Simulation

For each block, we simulate TeUS data using Field II [77] from the displacement data generated by the FEM. In US simulation, the speed of the
6.4 Experiments and Results

6.4.1 Feature Visualization Results

For the feature visualization experiment, by subtracting the distributions of GS pattern 3, 4 and benign samples in the input layer (Section 6.2), we found that feature one, corresponding to hidden activity of the first neuron of the third layer, along with features four and five, are those that maximally differentiate between different prostate cancer grades, especially in lower frequency range. Figure 6.5(a) shows the visualization of distribution differences for GS patterns 3, 4 and benign tissues related to the first learned
6.4. Experiments and Results

Figure 6.5: (a) Differences of distributions between cancerous tissues with Gleason patterns 3 and 4 as well as benign tissues back projected in the input neurons corresponds to the first neuron in the third hidden layer; (b) Spectral difference of the simulated TeUS in benign and different cancer tissues.

features of the third hidden layer, back propagated to the input layer. The results of feature visualization suggest that frequencies between $0 - 2$ Hz provide the most discriminative features for distinguishing between cancerous and non-cancerous tissues. The identified frequency range is also consistent with the ranges that we have observed in our previous independent in vivo studies [75, 76].

6.4.2 Simulation Results

Figure 6.5(b) depicts the spectral difference of the simulated TeUS using Field II, for GS3, GS4 and benign ROIs extracted from digital pathology slides. The simulation results also show a noticeable amplitude differences between benign and cancerous tissue regions in the lower frequency range.

Figure 6.6 depicts the distribution of the power spectrum of TeUS at 1 Hz excitation frequency (Fig. 6.6(a)) and its first harmonic (Fig. 6.6(b)) across benign, GS3 and GS4 ROIs. The distributions in both frequencies are statistically significantly different between benign and cancerous ROIs (all $p < 0.001$ using a paired t-test).
6.5. Conclusion

Evidence derived from our deep learning-based feature visualization pointed to low-frequency components of TeUS as the most informative features for tissue classification. These components potentially represent the effect of pulsation on prostate tissue microstructure. As a result, we simulated mechanical micro-vibrations of scatterers in phantoms with various scatterer distributions, reflecting benign and cancerous tissue, derived from digital histopathology data. We demonstrated that the micro-vibrations of scatterers could be captured by low-frequency spectral features of TeUS, similar to our in vivo results. These observations together with our previous results suggest that the distribution and micro-vibration of scatterers could lead to tissue typing information in TeUS.

Figure 6.6: (a) Distribution of the power spectrum in the frequency spectrum of simulated TeUS data at the excitation frequency, (b) Distribution of the power spectrum in the frequency spectrum of simulated TeUS data at the first harmonic of the excitation frequency.
Chapter 7

Conclusion and Future Work

It is difficult to say what is impossible, for the dream of yesterday is the hope of today and the reality of tomorrow.
— Robert H. Goddard

The ultimate diagnosis of prostate cancer is through histopathology analysis of prostate biopsy, guided by either Transrectal Ultrasound (TRUS), or fusion of TRUS with mp-MRI. One million patients in North America undergo TRUS-guided prostate biopsy annually. Among this cohort, 70% of 10 core biopsies return negative while up to 34% of the positive yield are undergraded. Computer-aided diagnosis models for detection of prostate cancer and guidance of biopsy involve both ultrasound and mp-MRI-based tissue characterization. mp-MRI has high sensitivity in detection of prostate lesions but low specificity, hence, limiting its utility in detecting disease progression over time. Ultrasound-based tissue characterization methods focus on the analysis of texture and spectral features within a single ultrasound frame, Doppler imaging, and elastography. Temporal Enhanced Ultrasound (TeUS), involving a time-series of ultrasound RF/B-mode frames captured from insonification of tissue over time, has enabled the depiction of patient-specific cancer likelihood maps. Despite promising results in detecting prostate cancer, accurate characterization of aggressive lesions from indolent ones is an open problem and requires refinement.

7.1 Conclusion and Summary

In this thesis, in an effort to improve TeUS-based tissue characterization, we developed methods and algorithms to automate diagnosis of prostate cancer and to accurately characterize aggressive lesions. A deep learning framework was devised, deployed and evaluated using TeUS data, to overlay near real-time cancer likelihood maps on TRUS images in a clinical environment. The proposed solution encapsulates variability associated with access to raw ultrasound signals in commercial scanners. This approach also provides complementary information about the grade and extent of prostate cancer.
7.1. Conclusion and Summary

The algorithms and methods were developed using data obtained through a study that was approved by the ethics review board of the National Cancer Institute (NCI), National Institutes of Health (NIH) in Bethesda, Maryland. In Chapter 2, the clinical process of TeUS data acquisition and establishing histopathology ground truth for the data is described. Through this thesis, the focus was on both time-domain and spectral-domain representation of TeUS signals as presented in Chapter 2.

In Chapter 3, frameworks were proposed for diagnosis of prostate cancer using TeUS. In this chapter, both the spectral and temporal aspect of TeUS data were modeled to find a more accurate technique for prostate cancer detection. The focus of the chapter was to distinguish between cancerous and normal tissue types represented as a binary classification task. A DBN-based approach was proposed to automatically analyze the spectral aspect of temporal ultrasound data obtained from 255 cancer foci identified in mp-MRI. Also, deep RNN were proposed to explicitly model the temporal information in TeUS. By investigating several RNN models, it was demonstrated that Long Short-Term Memory (LSTM) networks achieve the highest accuracy in detecting prostate cancer in TeUS data. Algorithms were presented for feature visualization using DBN and LSTM networks. The focus of In Chapter 4 was on detection of higher grade prostate cancer and the problem of separation between different prostate grades. An approach was presented to alleviate the challenge of noisy and sparse labels in building computer-aided diagnosis models for prostate cancer biopsy guidance. Towards realizing a practical solution for prostate cancer diagnosis using TeUS data in the clinical setting, in Chapter 5, the challenge of accessibility of raw RF data in commercial scanners was addressed. A method for knowledge transfer between TeUS RF and B-mode data was proposed. This RNN-based solution also enabled near real-time processing of RF/B-mode TeUS data. Finally, in Chapter 6, the physical phenomena underlying TeUS was investigated. In a hypothesis-generating study, to obtain insight into the physical phenomenon governing the interaction of temporal ultrasound with tissue, a deep learning based feature visualization method was proposed. Evidence derived from deep learning-based feature visualization pointed to low-frequency components of TeUS as the most informative features for tissue classification. Simulations were designed with mechanical micro-vibrations of scatterers in phantoms with various scatterer distributions, reflecting benign and cancerous tissue, derived from digital histopathology data.

In conclusion, in this thesis, the aim was to advance TeUS-based tissue characterization method and enable more precise prostate biopsy guidance
compared to the state of the art. The work aims to help increase the detection rates of prostate biopsy which ultimately, can lead to the reduction in the need for repeated biopsies. Results demonstrated that TeUS is effective in differentiating aggressive prostate cancer from clinically-less-significant disease and non-cancerous tissue. The proposed solutions have the potential to establish decision support models for patient-specific targeting during biopsy by displaying near real-time cancer likelihood maps on B-mode ultrasound images to indicate the presence of cancer.

The contributions of this thesis are summarized as follows:

- An automatic feature selection framework for spectral analysis of TeUS signal using DBN is proposed.

- Results demonstrated a statistically significant improvement in the accuracy of prostate cancer detection in compared to previously published studies using spectral features of TeUS signals [74, 76].

- A feature visualization approach is developed to determine the characteristics of non-cancerous and cancerous cores in TeUS data and their correlation with learned high-level features. This approach is used to identify the most discriminating frequency components of the time series as learned by the classifier.

- TeUS data is extensively and explicitly analyzed in temporal domain using time-dependent probabilistic deep networks for the first time.

- Results indicated that temporal analysis of TeUS using RNN can identify patterns in data that may not be readily observable in spectral domain analysis, leading to significant improvements in detection of prostate cancer.

- Algorithms were developed for in-depth analysis and visualization of high-level latent features of LSTM-based RNN.

- A transformational finding, achieved through this analysis, is that the most discriminating features for detection of prostate cancer can be learned from a fraction of the full TeUS time series. This information can be used to optimize TeUS data acquisition for clinical translation.

- A novel approach for grading and detection of high-grade prostate cancer using spectral analysis of TeUS with DBN is proposed.
7.1. Conclusion and Summary

- This approach could successfully differentiate among aggressive prostate cancer (GS\(\geq 4+3\)), clinically less significant prostate cancer (GS\(\leq 3+4\)), and non-cancerous prostate tissues.

- A novel approach is devised for grading and detection of high-grade prostate cancer using temporal analysis of TeUS with RNN. By encapsulating proposed ground-truth probability vectors, this solution can also precisely estimate cancer length in biopsy cores.

- The accuracy of tissue characterization is statistically significantly improved as compared to previously published studies [10, 12].

- A novel strategy is proposed for depiction of patient-specific colormaps for biopsy guidance including the estimated model uncertainty. The method could highlight the possible misguidance in biopsy by using the uncertainty measure.

- To address limited access to raw ultrasound data on commercial scanners, a transfer-learning strategy is developed to enable the probabilistic modeling of TeUS B-mode data.

- The viability of using B-mode TeUS for cancer detection is demonstrated using retrospective data. The initial evaluation indicates that the solution is capable of providing guidance information for prostate biopsy procedures.

- Near real-time augmentation of live standard 2D ultrasound images with cancer likelihood maps generated from the models is implemented.

- A method for interpretation and visualization of the high-level learned features from TeUS data is presented. Evidence derived from feature visualization points to low-frequency components of TeUS as the most informative features for tissue classification. These components potentially represent the effect of pulsation on prostate tissue microstructure in form of micro-vibrations.

- The effect of micro-vibration is simulated using a medium with preset elasticity and scatterer locations extracted from 14 whole-mount digital histopathology slides.

- Results showed that the distribution and micro-vibration of scatterers could lead to tissue typing information in TeUS. This finding is a major breakthrough in understanding and technical formulation of TeUS after a decade.
7.2 Future Work and Suggestions for Further Development

A number of interesting areas of research can be suggested to further optimize and improve the current solutions as follows:

- The immediate envisioned future direction for the current thesis is a large inter-institution patient study to determine the accuracy of the proposed prostate cancer grading and diagnosis approaches across a wide range of patient sub-populations. Access to a larger dataset would also allow to further explore machine learning solutions to address the issue of noisy and sparse labels. These include:
  - Adding an extra noise layer to the network which adapts the network outputs to match the noisy label distribution. The parameters of this noise layer can be estimated as part of the training process and involve simple modifications to current training infrastructures for deep networks [145]. Direct modeling of the noise corresponding to the label and the ultrasound data may require a small data set with more reliable labels [47].
  - Developing multi-instance learning approaches which can combine probabilistic modeling at ROI and core levels in the same solution [157]. For this purpose, the use of multi-instance neural networks is proposed which perform multi-instance learning in an end-to-end manner. These solutions takes bags (here cores) with a various number of instances (here ROIs) as input and directly output the labels of the bags.

- A fundamental assumption in previous models for temporal enhanced ultrasound is that data are collected in one center with specific equipment and image settings. However, center-specific differences in data exist in form of variations in the imaging equipment, clinician preferences for image parameter settings, and the patient pool. Hence, it may not be possible to readily use a model built using data from one center in other centers. It is recommended to use an automatic transfer learning between different image settings and different centers. The solution can learn discriminant features from ultrasound signals in the source dataset and use a transfer learning method, similar to what we discussed in Chapter 5 to take into account the center-specific differences.
7.2. Future Work and Suggestions for Further Development

- One of the limitations of the current study is that all of the cores come from prostate regions where the presence of cancer is suspected in mp-MRI. Future work should be focused on the evaluation of the method for the cases that mp-MRI does not show any suspicious cancer region. This can be achieved as part of systematic prostate biopsy studies.

- There has been limited research over the past few years to combine Bayesian models and deep networks for estimation of uncertainty in decision-making models. However, these preliminary methods have limitations including severe underestimate of model uncertainty [48, 93]. Exploring this path and designing probabilistic models that determine “How uncertain are we about the decision of the model?” is recommended. This concept has an special importance for any computer-aided diagnosis system. Also, future work can be focused on the analysis of the source of the uncertainty and integrate the proper solution in the framework.
Bibliography


[36] Mohammad I Daoud, Parvin Mousavi, Farhad Imani, Robert Rohling, and Purang Abolmaesumi. Tissue classification using ultrasound-


Farhad Imani, Purang Abolmaesumi, Eli Gibson, Amir Khojaste, Mena Gaed, Madeleine Moussa, Jose A Gomez, Cesare Romagnoli, D Robert


[120] Kathryn R Nightingale, Mark L Palmeri, Roger W Nightingale, and Gregg E Trahey. On the feasibility of remote palpation using acoustic


Bibliography


[159] Shao Wei Xie, Hong Li Li, Jing Du, Jian Guo Xia, Yi Fen Guo, Mei Xin, and Feng Hua Li. Influence of serum PSA level, prostate volume, and PSA density on prostate cancer detection with contrast-enhanced sonography using contrast-tuned imaging technology. Journal of Ultrasound in Medicine, 32(5):741–748, 2013.


Appendix A

Theoretical Background of Temporal Enhanced Ultrasound

The characteristic model for the formulation of RF backscattered ultrasound signal can be expressed as [104, 127]

\[ I(x) = P(x) \ast S(x) + n, \]  

(A.1)

where \( P(x) \), ultrasound Point Spread Function (PSF) at an arbitrary point \( x \), is convolved with \( S(x) \), the tissue scattering function at the same point, to generate \( I(x) \), the backscattered RF data at point \( x \). In this equation, \( n \) represents random noise.

If we assume that \( (S(x)) \) varies locally as a function of time, a simple model of this effect at point \( x_0 \) can be expressed as \( S(x_0 + f(t)) \), where \( f(t) \) is a time varying function and \( t \) is the “slow time” (i.e., frame number). For relatively small \( f(t) \), the first order approximation of Taylor expansion of \( S \) can be written as

\[ S(x_0 + f(t)) = S(x_0) + \frac{\partial S(x)}{\partial x} \bigg|_{x=x_0} f(t), \]  

(A.2)

where \( \frac{\partial S}{\partial x}(x_0) \) is the local change in the scattering function about point \( x_0 \). As stated in the Chapter 6, we previously investigated a scenario where variations in \( f(t) \) were primarily due to induced thermal effects as a result of acoustic absorption [36]. In this thesis, we investigate an alternative hypothesis, where variations in \( f(t) \) are a result of tissue micro vibrations due to physiological vibration. Let \( f(t) \) be a sinusoidal function of the form:

\[ f(t) = a(x_0) \sin(\omega t), \]  

(A.3)

where \( a(x_0) \) is the amplitude of the micro vibration at point \( x_0 \), and \( \omega \) is the frequency of vibration. In a fully elastic tissue, \( a(x_0) \) is inversely proportional
Appendix A. Theoretical Background of Temporal Enhanced Ultrasound

to $E(\omega)$, the local elasticity, which is frequency dependent. In a viscoelastic medium, $a(x_0)$ is inversely proportional to a function of $E(\omega)$ and $\mu(\omega)$, where $\mu(\omega)$ is the local viscosity, which is also frequency dependent [162]. Equation (A.2) can be rewritten as

$$S(x_0 + a(x_0) \sin(\omega t)) = S(x_0) + a(x_0) \left. \frac{\partial S(x)}{\partial x} \right|_{x=x_0} \sin(\omega t).$$  \hspace{1cm} (A.4)

Combining Equations (A.1) and (A.4), we have:

$$I(x_0, t) = (P(x) * S(x)) \bigg|_{x=x_0} + a(x_0) \left. \left( P(x) * \frac{\partial S(x)}{\partial x} \right) \right|_{x=x_0} \sin(\omega t) + n. \hspace{1cm} (A.5)$$

The first term of Equation (A.5) corresponds to the time-invariant component of the RF signal received from point $x_0$. This component depends only on the spatial variations of the backscattering function across the propagation medium. It can be characterized using conventional analysis of the RF spectrum and B-mode texture. The second term corresponds to time-varying components of $I(x_0, t)$ affected by local variations in the backscattering function, in slow time. Such local variations represent changes in tissue structure, such as changes in nuclear configuration.

Two important observations can be made about the second term of this equation: 1) In media with the same mechanical properties $E$ and $\mu$, spectral analysis of TeUS captures $(P(x) * \frac{\partial S(x)}{\partial x}) \bigg|_{x=x_0}$, which is related to spatial variations in the scattering function. This property can be of benefit to characterize, e.g., tissue at early stage cancer, where changes in nuclei configuration could dominate changes in tissue property; 2) Where there are changes in mechanical properties, $a(x_0)$, and the scattering function, $S$, or the vibration frequency, $\omega$, TeUS captures a combined effect for tissue characterization.

Through a series of simulations [19], we demonstrate that local changes in tissue properties, captured by $a(x_0) \left. (P(x) * \frac{\partial S(x)}{\partial x}) \right|_{x=x_0}$ using, e.g., Fourier transform of TeUS, are effective features for tissue characterization. Our simulations include media with synthetic array of scatterers of varying distances and pathology mimicking simulations based on whole-mount prostate
Appendix A. Theoretical Background of Temporal Enhanced Ultrasound

digital histopathology data. These simulations confirm our observations from Equation (A.5) that as a result of micro vibration in the medium, TeUS can differentiate tissue types with subtle variations in arrangement of scatterers.
Appendix B

TeUS Biopsy Guidance System Implementation

In this part, we present the details of our unified software framework demonstrating real-time analysis of ultrasound data stream using a deep learning solution. The system integrates cutting-edge machine learning software libraries with 3D Slicer [41] and the “Public software Library for UltraSound imaging research (PLUS)” [91] to build an accessible platform. To the best of our knowledge, this is the first system of its kind in the literature. This is the very first demonstration of automatic, real-time prostate segmentation in ultrasound in the literature. The proposed software system allows for depiction of live 2D US images augmented with patient-specific cancer likelihood maps that have been calculated from TeUS.

B.1 TeUS biopsy guidance system

The overview of the components of the guidance system is given in Fig. B.1 and Fig. B.2. To allow for continuous localization of likely-cancerous tissue in US data, the architecture incorporates state-of-the-art open-source software libraries for US data streaming, data visualization, and deep learning. A client-server approach allows running computationally expensive algorithms simultaneously, and in real-time. The system has a three-tiered architecture as seen in Fig. B.2: US-machine layer, TeUS-client, and TeUS-server. US-machine layer acquires data and streams it to the TeUS-client layer. TeUS-client is a 3D Slicer [41] extension responsible for US data management, pre-processing and visualization. Ultrasound B-mode data, or B-mode and Radio Frequency (RF) data if both available, are streamed to the TeUS-server for tissue characterization and prostate localization, and received back by the TeUS-client for real-time displaying in 3D Slicer. The TeUS-server receives the US data through an OpenIGTLink network protocol [149], performs segmentation for continuous tracking and localization of the prostate, and computes cancer likelihood maps that are transferred back to the TeUS-client.
B.1. TeUS biopsy guidance system

Figure B.1: Overview of the biopsy guidance system. The three steps in the guidance workflow are volume acquisition, classification and guidance. A client-server approach allows for simultaneous and real-time execution of computationally expensive algorithms including TeUS data classification, and prostate boundary segmentation.

through a second OpenIGTLink. The following sub-sections describe details of TeUS-client and TeUS-server.

B.1.1 TeUS-client

The TeUS-client’s primary tasks are receiving streamed B-mode data and generating a TeUS sequence, preprocessing the data to divide it to smaller Regions of Interest (ROIs), and visualizing the cancer likelihood maps overlaid on US data. Data from ROIs are sent to the TeUS-server for analysis through the machine learning framework. TeUS-client receives the results as a colormap and segmented prostate boundary, and overlays this information on the US image. The TeUS-client includes a custom $C++$ loadable module, TeUS guidance module, created as an extension to 3D Slicer (Fig. B.2).

Ultrasound data acquisition: Once the TeUS guidance module is started,
B.1. TeUS biopsy guidance system

Figure B.2: The software system has a three-tiered architecture. Ovals represent processing elements while arrows show the direction of data flow. In the US machine layer, PLUS is responsible for US data acquisition and communicates with the TeUS-client via the OpenIGTLink protocol. The TeUS client layer includes TeUS guidance, an extension module within the 3D Slicer framework. The TeUS-server layer is responsible for the simultaneous and real-time execution of computationally expensive algorithms and communicates with TeUS-client via the OpenIGTLink protocol.

an instance of PLUS [91] is initiated on the US-machine layer by running the PlusServer application. Next, an OpenIGTLink receiver is created that continuously listens to the PlusServer to receive the US data (Fig. B.2). Upon successful connection, the US data are displayed on the 3D Slicer window. Once the user issues a “Start” signal, data acquisition is initiated by buffering streamed B-mode images, followed by preprocessing. Data acquisition is halted by the user through a “Stop” signal.

**Preprocessing:** Once the data acquisition begins, a independent preprocessing thread is initiated to avoid freezing 3D Slicer while the TeUS ROIs are being generated. In this thread, each US image is divided to equally-sized ROIs of 0.5 mm × 0.5 mm based on the scan conversion parameters. For the \(i^{th}\) ROI, a sequence of TeUS data, \(x^{(i)} = (x_1^{(i)}, ..., x_T^{(i)})\), \(T = 100\) frames, is generated by averaging all time series within that ROI and subtracting the mean value from the given time series.

**Communication with the TeUS-server:** Following the completion of the preprocessing thread, the TeUS guidance module sends data from the extracted ROIs to the TeUS-server layer, using a standard OpenIGTLink
B.1. TeUS biopsy guidance system

protocol and by forking a sender thread. In addition to this thread, the TeUS-client layer also forks a receiver thread, which creates a new connection with the TeUS-server using the OpenIGTLink protocol. The receiver thread then waits for a message containing the resulting cancer likelihood colormap as well as the prostate localization information from the TeUS-server.

Receiving and visualizing the cancer likelihood colormap: Upon the generation of results and receiving the output message from the TeUS-server, the TeUS-client’s receiver thread picks up the TeUS module’s execution. The guidance colormap and prostate boundary segmentation results are saved and 3D Slicer thread begins overlaying the information. During the guidance colormap visualization, the segmentation information is used to localize the prostate and mask out any colormap data falling outside the boundary. The boundary matrix is resized down to the colormap’s dimensions. Then, the guidance colormap is converted from single channel float values (ranging from 0 to 1) to 3-channel RGB data (ranging from 0 to 255), with 0 being pure Blue and 1 being pure Red (The Green channel is always 0). The boundary mask and processed colormap are multiplied together, masking out any data outside the prostate boundary.

B.1.2 TeUS Server

The integrated system encapsulates a machine learning framework where we specifically use two deep learning methods, implemented in Tensorflow [1]. These deep networks are responsible to identify target locations using TeUS data and a state-of-the-art automatic prostate boundary segmentation technique is used to localize the prostate boundary during the prostate biopsy guidance. The TeUS-server is mainly responsible for simultaneous and real-time execution of computationally expensive deep learning models. The TeUS-server loads, prepares and runs a Tensorflow graph [1] from a saved protocol buffers (.pb) file. Our Tensorflow graphs include the trained model parameters obtained from deep learning methods we will explain below. The TeUS-server also receives the extracted TeUS ROIs from the TeUS-client, buffers them into a Tensorflow Tensor object, and after running the Tensorflow graph, returns the result back to the TeUS-client. The TeUS-server is a standalone C++ application running on Linux, and is built from within a clone of Tensorflow and is compiled using Bazel open-source toolbox.

Receiving the TeUS data: As with the TeUS-client, the TeUS-server creates two socket-based objects: a sender and a receiver. Note that the TeUS-server’s receiver receives the TeUS ROIs data from the client’s sender,
while the TeUS-server’s sender sends the output colormap to the client’s receiver. There is no need for multi-threading on the TeUS-server side because it needs to receive the frames in sequential order.

**Running the deep neural networks:** Concurrent with the OpenIGTLink activity, the TeUS-server also performs a few steps to set up and run the Tensorflow graphs for PCa detection and prostate boundary segmentation. The TeUS-server initializes the Tensorflow graph, by loading the cancer classification and segmentation trained models from the protocol buffer files and initializes the buffer tensor, which will be filled with $T = 100$ frames. After receiving the $T^{th}$ frame, the TeUS-server runs the Tensorflow session, feeding the buffer tensor as input to the classification network graph. Simultaneously, prostate boundary segmentation graph is fed with the buffer tensor as the input to generate the segmentation results. Then, TeUS-server converts the returned tensors into a float vector, packs the float data from output vectors into a new OpenIGTLink “ImageMessage” and sends them to the TeUS-client using the TeUS-server’s sender object.

**Prostate cancer classification:** The deep networks are generated based on the methods that we presented in our earlier works and from a data set consisting of biopsy targets in mp-MRI-TRUS fusion-biopsies with 255 biopsy cores from 157 subjects (here, we refer to this data as the first retrospective study). We give a brief overview of these methods. For a detailed description of the models, the reader may refer to [10, 14].

**Prostate segmentation:** The segmentation deep networks are generated based on our earlier works [5]. The network is pre-trained on a dataset consisting of 4,284 expert-labeled TRUS images of the prostate as explained in [5] and further fine-tuned using manually segmented B-mode images obtained during mp-MRI-TRUS fusion biopsy from $D_{train}$ subjects. The method is based on residual neural networks and dilated convolution at deeper layers. The model takes an input image of the size of $224 \times 224$ pixels and generates a corresponding label map of the same size. For a detailed description of the models, the reader may refer to [5].