Test-Time Training for Prostate Cancer Detection: Addressing Distribution Shift with Self-Supervised Learning

Mahdi Gilany[1], Paul Wilson[1], Amoon Jamzad[1], Mohammed Harmanani[1], M. N. Nhat To[2], Brian Wodlinger[3], Purang Abolmaesumi[2], Parvin Mousavi[1]

[1]Queen's University, Kingston, Ontario, Canada, [2] University of British Columbia, Vancouver, British Columbia, Canada, [3]Exact Imaging, Markham, Ontario, Canada

INTRODUCTION: Histopathological examination of prostate tissue samples, typically obtained through transrectal ultrasound (TRUS)-guided biopsies, remains the prevalent method for diagnosing and grading prostate cancer (PCa). However, TRUS approach is often hindered by a high false-negative rate and substantial risks of negative biopsy-related effects. To enhance the PCa detection via ultrasound, a range of deep learning techniques have been introduced. These methods utilize ultrasound Bmode or radio-frequency images to determine the characteristics of the underlying tissues, thereby pinpointing areas that may require closer examination. Nonetheless, a significant challenge with these deep learning techniques is the data distribution shift across different hospitals, patients, or even within multiple images from the same patient, which is a crucial factor for the real-world deployment of these models. To overcome this obstacle, we suggest a strategy of adapting the deep model to the distribution of the test data during the inference phase through the use of self-supervised learning.

METHODS: Our study utilizes a proprietary dataset obtained using advanced micro-ultrasound technology (29 MHz center frequency) from 693 patients across five centers. These patients were part of a clinical trial (NCT02079025) and had undergone systematic TRUS-guided biopsies. Our approach tailors test-time training [1] to adapt the model to the test distribution. As depicted in Figure 1, we employ ResNet10 as the feature extractor to derive feature vectors from each region of interest (ROI) in both the prostate area (unlabeled data) and the needle region (labeled data). The classification head and self-supervised learning (SSL) heads are employed to transform the feature vectors derived from labeled and unlabeled regions of interest (ROIs) into classes and SSL features, respectively. We utilize cross-entropy for the classification head, while the SSL head is trained using "Bootstrap Your Own Latent (BYOL)" method [2] loss. During the training phase, the feature extractor and both heads are concurrently trained. In the inference phase, the feature extractor and SSL head are fine-tuned using unlabeled test data, adjusting to test distribution.

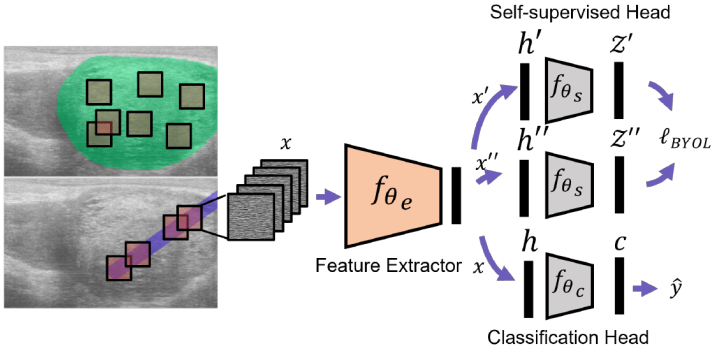

Figure 1: Summary of the proposed method: unlabeled ROI patches are extracted from prostate region to train feature extractor and SSL head. Labeled ROI patches are extracted from needle region where pathology annotation exist, and they train feature extractor and classification head. During inference, unlabeled test data is used to fine-tune the feature extractor before prediction is made.

RESULT & CONCLUSION: Our proposed method, in Table 1, has significantly improved performance compared to the baseline approach, as evidenced by the 4% improvement in ROI-wise AUROC and 5% in core-wise AUROC. Test-time training, applied to test data prior to making predictions, enables the model to adapt to the test distribution, thus significantly enhancing its real-world robustness and performance.

Table 1: Comparison of the proposed method to baseline methods. Metrics are threshold independent area under the ROC curve (AUROC) for both ROIs and cores (all ROIs in the needle).

Method	Core AUROC	ROI AUROC
ResNet10 [3]	71.1	62.1
TTT-ResNet10 (ours)	**76.6**	**66.6**

REFERENCES: [1] Grill et al., Neurips 2020 [2] Sun et al., ICML2020 [3] Gilany et al., MICCAI 2022