

ProstNFound-RL: Guided Attention with Reinforcement Learning for Robust Prostate Cancer Detection

Mohammad Mahdi Abootorabi
`mahdi.abootorabi@ece.ubc.ca`

1 Objective / Goal

Introduction and Related Works. Prostate cancer (PCa) is one of the leading causes of cancer-related death among men, underscoring the need for early and accurate detection. The current diagnostic standard, transrectal ultrasound (TRUS)-guided biopsy, suffers from the poor visibility of lesions in conventional ultrasound, resulting in random sampling that may miss clinically significant tumors or overdiagnose insignificant ones. Recent advances in micro-ultrasound combined with AI have shown promise in improving detection accuracy. ProstNFound [1] leverages medical foundation models enhanced with ultrasound domain knowledge and clinical biomarkers for robust PCa detection. Cinepro [2] applied robust training and temporal data augmentation on ultrasound cineloops to mitigate label noise. ProTeUS [3] further integrates spatio-temporal features from time-series RF ultrasound signals and clinical data, using a progressive training strategy to achieve a great performance. Collectively, these models demonstrate the potential of adapting foundation models with domain-specific information to improve prostate cancer diagnosis from ultrasound imaging.

Despite this progress, significant challenges remain. Supervision is typically weak, limited to noisy labels at the needle region rather than precise pixel-level annotations. Available labels often provide only cancer involvement scores rather than exact tumor boundaries. Furthermore, the scarcity of training data hampers the development of AI systems that generalize well while maintaining high diagnostic accuracy.

The integration of Reinforcement Learning (RL) into this domain remains largely unexplored. While existing models perform well as static risk calculators, they leave the crucial task of visual search and evidence integration to the human operator. This project aims to overcome that limitation by introducing a novel RL-guided attention mechanism. The core idea is to train an RL agent on top of existing foundation-based models to actively explore ultrasound images, identifying suspicious regions and enhancing both detection performance and clinical interpretability. We hypothesize that RL is particularly well-suited for this problem because it enables active search policies, learning where to look, to mimic expert decision-making during ultrasound scanning. Initially, we plan to implement this using Group Relative Policy Optimization (GRPO) [4]. However, we may later explore other modern RL algorithms that better align with the problem’s constraints and data limitations.

2 Methodology / Experiments

We will build upon the existing ProstNFound+ architecture, which employs the powerful MedSAM [5] foundation model as its backbone and extends it with clinical awareness for prostate cancer detection. Our methodology integrates a Reinforcement Learning (RL) agent that learns to guide the focus of the MedSAM [5] decoders—adapted in ProstNFound+ for heatmap and risk prediction tasks. Figure 1 illustrates the architecture of ProstNFound+, showing the encoding–decoding process that currently achieves state-of-the-art performance in this field.

Our approach trains an RL agent to act as an intelligent focus mechanism. The agent learns a policy for identifying the most suspicious regions within the prostate image, providing critical contextual evidence to the decoders before they make a final prediction for a given biopsy core.

Proposed Architecture. (i) We use the frozen MedSAM [5] image encoder from ProstNFound+ to generate a rich feature map for each ultrasound image. This global representation serves as the state for our RL agent. (ii) A lightweight, trainable policy network is implemented to analyze the encoder’s feature map and identify the k most suspicious regions that warrant closer inspection. The agent’s action is to output the coordinates of these regions. (iii) These coordinates are encoded as attention prompts in addition to clinical data, instructing the decoders to focus on specific image regions likely to contain pathological cues. This guidance

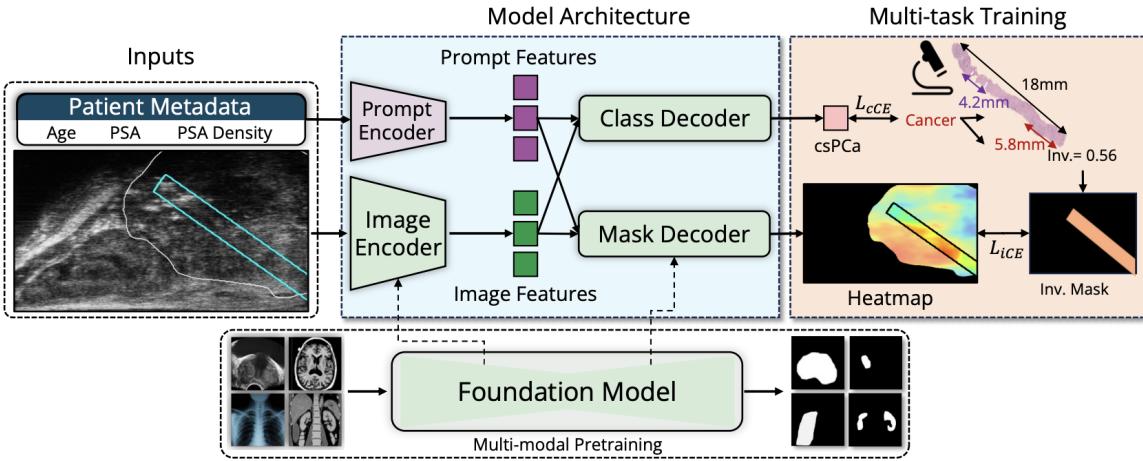


Figure 1: ProstNFound+ Architecture

enhances the decoders’ contextual awareness, potentially improving diagnostic accuracy. **(iv)** The system is trained end-to-end. The RL agent receives a reward when the regions it identifies help the decoders make correct predictions. Simultaneously, the ProstNFound+ decoders are fine-tuned to effectively utilize the agent’s guidance. The reward signal is designed to emphasize clinical significance, providing higher rewards for correctly identifying clinically significant prostate cancer (csPCa) and penalizing incorrect or irrelevant detections. Initially, the RL agent will be trained using Group Relative Policy Optimization (GRPO) [4]. However, other modern RL algorithms may later be explored to address domain-specific limitations and optimize performance.

Dataset. We will use the multi-center micro-ultrasound (μ US) dataset employed in previous studies. Each biopsy core is represented by one sagittal B-mode μ US image with its corresponding annotated needle trace. The dataset includes approximately 750 patients and 7,500 biopsy cores, providing a diverse set of examples for training and evaluation.

Evaluation. We will evaluate our method using the same quantitative metrics as prior work, including accuracy, AUROC, and sensitivity, for both classification and cancer heatmap generation tasks. In addition, we will conduct a qualitative evaluation of the RL agent’s behavior, analyzing whether the identified attention points align with clinically suspicious regions and how they influence the overall cancer detection performance. For baselines, we will compare our RL-augmented model against the original ProstNFound+ architecture (without RL). If results are promising, we will further compare against other state-of-the-art baselines, such as Cinepro [2].

3 Expected Results

We anticipate that ProstNFound-RL will improve diagnostic accuracy over ProstNFound+. Quantitatively, we expect higher AUROC and sensitivity for csPCa detection. Qualitatively, we expect the agent’s visualized attention points to correspond with clinically meaningful regions, demonstrating a learned and interpretable search policy. We also hypothesize that increasing the number of attention “glimpses” will further enhance performance, validating the effectiveness of the RL-guided attention mechanism.

4 Deliverables

The deliverables for this project include the implementation code and experimental results of the proposed ProstNFound-RL model. We acknowledge that this is an ambitious and high-risk idea, but its potential impact is substantial; it could lead to a novel, clinically applicable AI framework that meaningfully improves biopsy guidance in prostate cancer diagnosis.

We also recognize that the scope of this work exceeds that of a typical course project. Therefore, we plan to establish clear milestones to ensure steady progress and meet course requirements. The primary milestone is to implement the RL-guided focus mechanism and compare its performance against the baseline ProstNFound+ model. Depending on the outcomes, we aim to refine the approach further and potentially develop this work into a research publication for a top-tier AI conference. Although challenges are expected, due to the complexity of the base model, data characteristics, and training stability of RL, the project offers an exciting opportunity for innovation and real-world clinical impact.

References

- [1] P. F. Wilson et al., “Prostnfound: Integrating foundation models with ultrasound domain knowledge and clinical context for robust prostate cancer detection,” in *International Conference on Medical Image Computing and Computer-Assisted Intervention*, Springer, 2024, pp. 499–509.
- [2] M. Harmanani et al., “Cinepro: Robust training of foundation models for cancer detection in prostate ultrasound cineloops,” in *2025 IEEE 22nd International Symposium on Biomedical Imaging (ISBI)*, IEEE, 2025, pp. 1–5.
- [3] T. Elghareb et al., “Proteus: A spatio-temporal enhanced ultrasound-based framework for prostate cancer detection,” in *International Conference on Medical Image Computing and Computer-Assisted Intervention*, Springer, 2025, pp. 405–414.
- [4] Z. Shao et al., *Deepseekmath: Pushing the limits of mathematical reasoning in open language models*, 2024. arXiv: [2402.03300 \[cs.CL\]](https://arxiv.org/abs/2402.03300). [Online]. Available: <https://arxiv.org/abs/2402.03300>.
- [5] J. Ma, Y. He, F. Li, L. Han, C. You, and B. Wang, “Segment anything in medical images,” *Nature Communications*, vol. 15, p. 654, 2024.