Localizing Regions of Interest in Whole Slide Images via Reinforcement Learning

Harvard Medical School
Brigham and Women's Hospital, Founding Member, Mass General Brigham
Massachusetts General Hospital, Founding Member, Mass General Brigham

Bowen Chen, Richard Chen, Ming Y. Lu, Tiffany Y. Chen, Sharifa Sahai, Astrid Weins, Faisal Mahmood

Lab Website
@AI4Pathology
Mahmood Lab AI for Pathology
BROAD INSTITUTE

Abstract

- Growth of artificial intelligence (AI) has yielded many breakthroughs, from disease diagnosis to prognosis, and contributes to reducing variability in pathologist diagnoses.
- The most powerful of these AI models, however, depend on processing scanned images of whole pathology slides, which serve as the standard basis of pathology analysis.
- Whole slide images (WSIs) can range up to *gigapixels* in size, significantly larger than traditional computer vision applications. Thus, the deployment and scaling of such models require costly computing resources, which are infeasible in low resource settings that would most benefit from AI solutions.
- In this project, we present a reinforcement learning (RL)-based model to identify regions of interest (ROIs) in WSIs by selectively zooming into relevant patches of the image.
- We applied our model to the task of localizing glomeruli in kidney biopsies and found that our solution reduces the number of total pixels examined in the entire WSI by 70-fold and the tissue-containing region by 5-fold

2-Stage-Zoom Example Policies

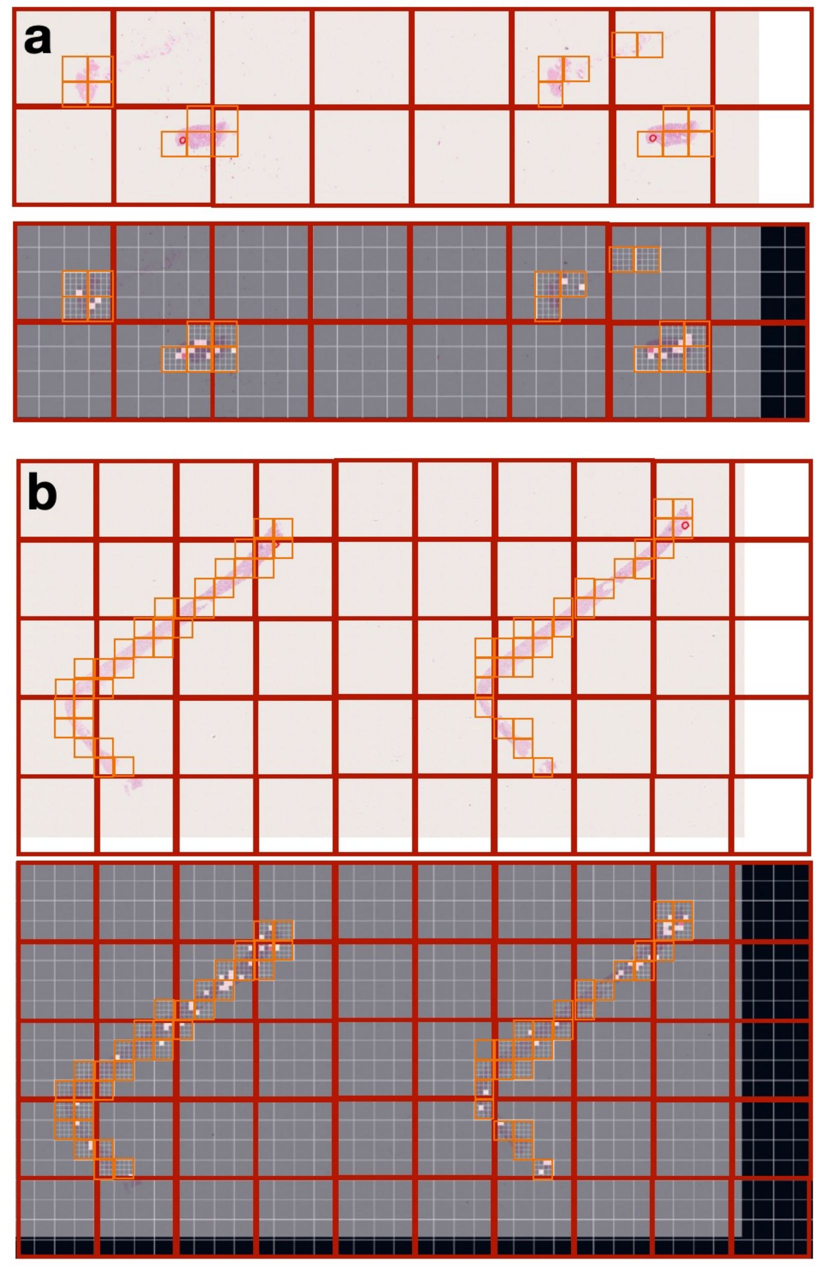

Fig 2.1. Selection policies for 2 examples WSIs output. For both **a** and **b**, top: policy overlaid on ground truth WSI. Policy overlaid on model predictions.

Red grid: all patches of WSI are seen at 0.625×.
Orange grid: patches selected at 2.5×. If not covered by grid, then not selected
Red patches: patches selected at 10× and classified as glomeruli
Light patches: patches selected at 10× and classified as non-glomeruli
Dark patches: patches not selected at 10×

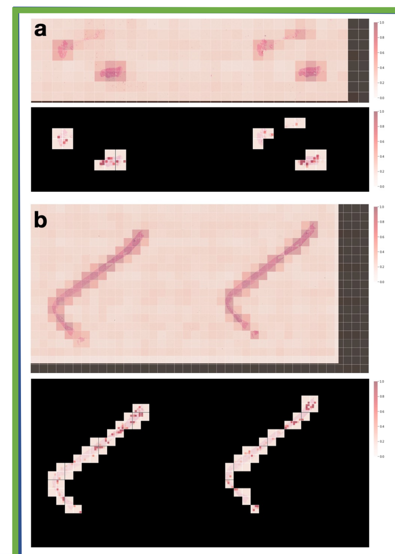

Fig 2.2. Selection probability scores from model for the same two examples as Fig 2.1

2-Stage-Zoom: Model Architecture and Workflow

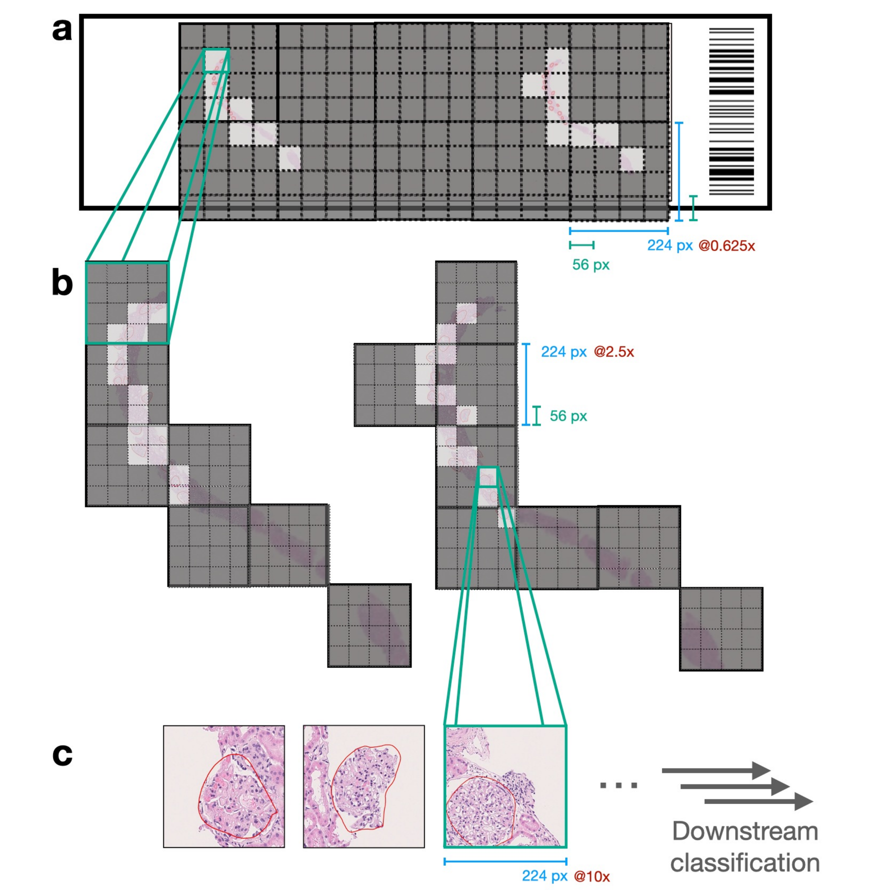

Fig 1.1. Ideal selection process with two zoom stages: 0.625× to 2.5× to 10×. Dark gray patches indicate not selected at that zoom stage; lighter patches indicate selected.

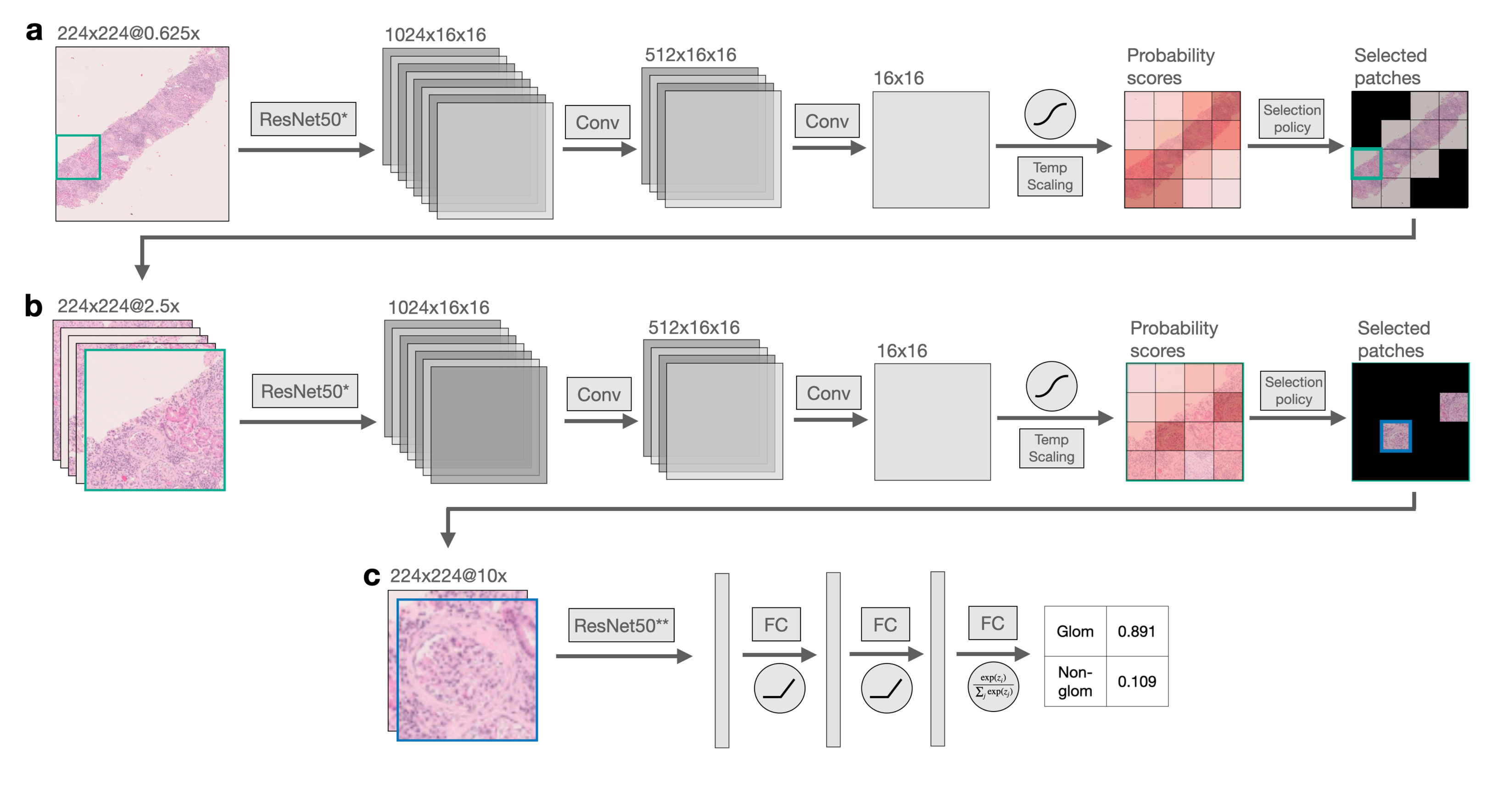

Fig 1.2. Example model workflow on one patch from a WSI. **a:** First zoom stage selects patches to zoom into for the second stage. **b:** Second zoom stage selects patches to further zoom into for classification. **c:** Selected zoom patches are fed into a classification network, which outputs probability scores for each patch.

Comparison with Naive and Alternative Policies

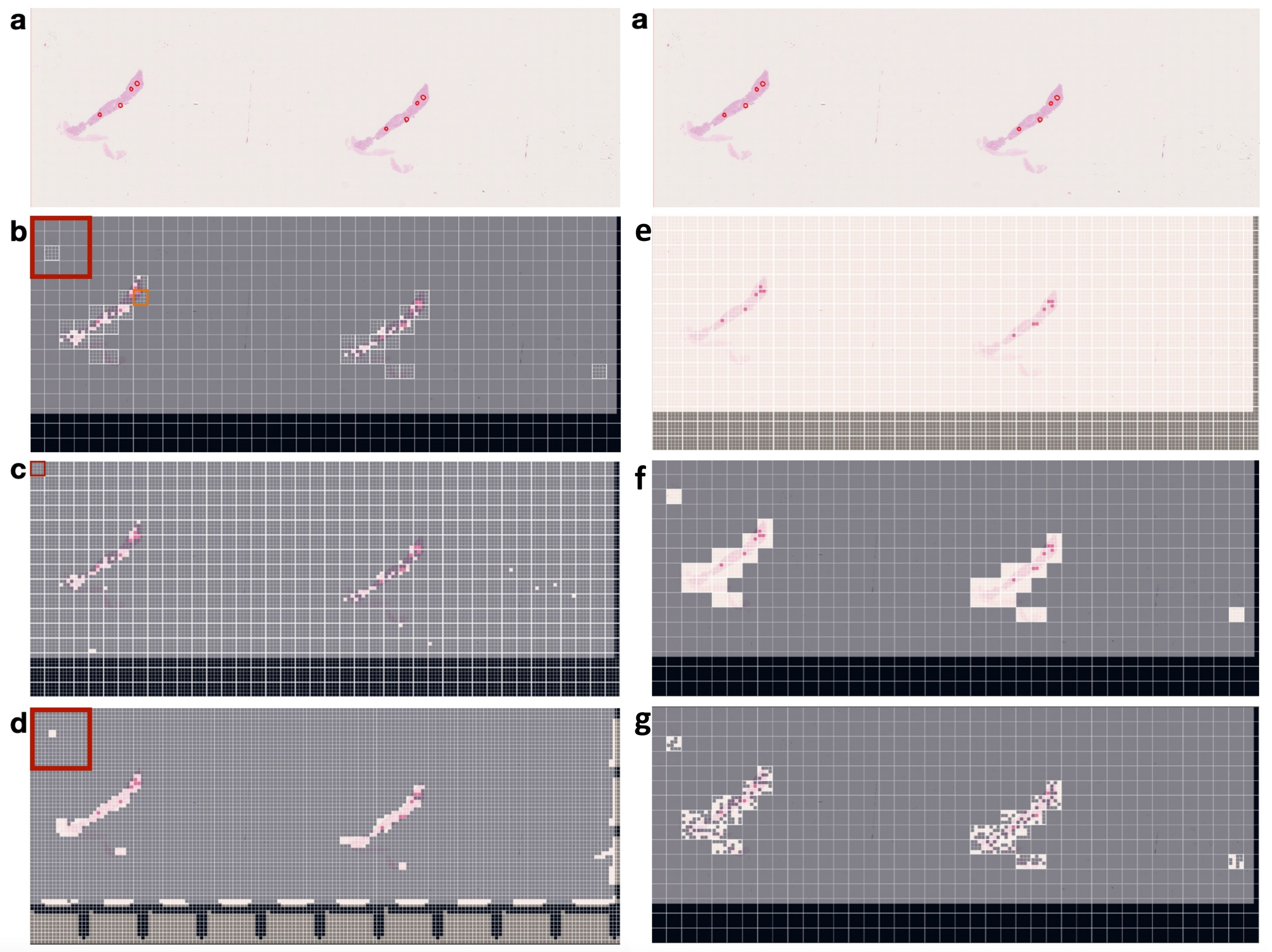

Fig 3.1. Example policies from comparison metrics on the same WSI. **a.** Ground truth. **b-d.** RL policies: 2-stage (0.625 to 2.5 to 10), 1-stage-4 (2.5 to 10), 1-stage-16 (0.625 to 10). **e-g.** Naive policies: selecting the entire slide, selecting all tissue, selecting random patches of tissue
Light patches: patches selected at 10× and classified as non-glomeruli
Dark patches: patches not selected at 10×

Model or naive policy	Glomeruli coverage	% Slide selected (@2.5×, @10×)	Pixels seen	% Glom correct given selected	% Glom selected and correct
2-stage	**96.0%**	**(6.14%, 0.613%)**	**7.41M**	**82.0%**	**77.9%**
1-stage-4	98.2%	(N/A, 0.695%)	37.1M	80.1%	78.3%
1-stage-16	93.0%	(N/A, 14%)	77.4M	78.3%	71.3%
All slide	100%	(N/A, 100%)	536M	78.7%	78.7%
All tissue	100%	(N/A, 6.16%)	34.9M	78.7%	78.7%
Random tissue	63.3%	(N/A, 3.02%)	18.2M	80.1%	43.4%

Fig 3.2. Summary table of comparison metrics. 2-stage-zoom drastically reduces pixels seen without loss of coverage or correctness

Agent Behavior

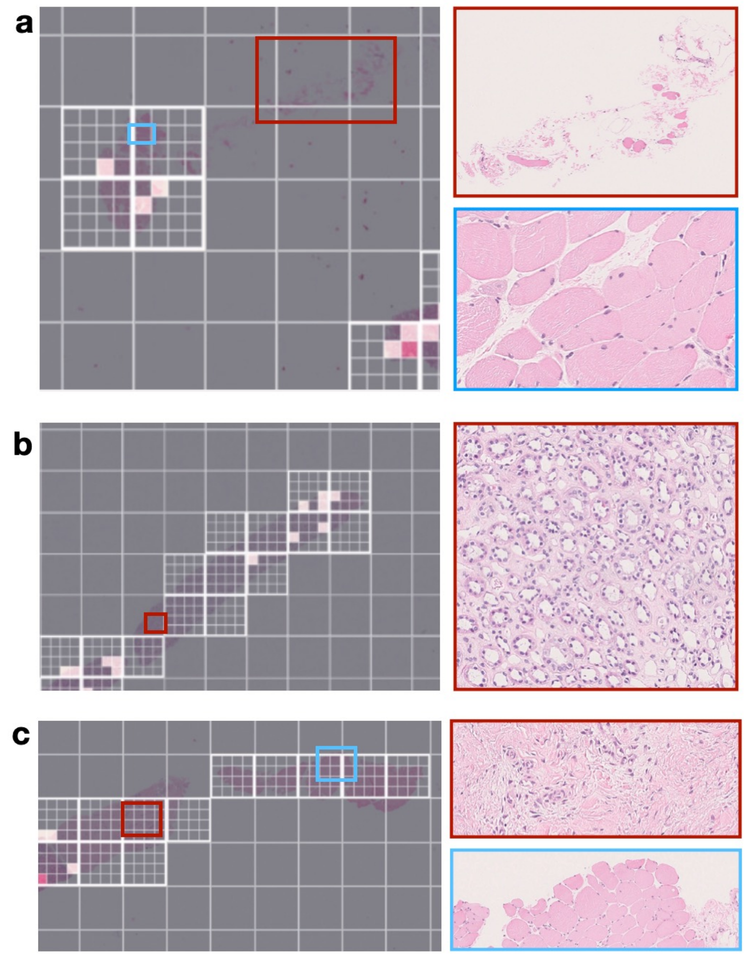

Fig 4.1. Interpretation of selection policy based on tissue type.
a. Red: fragments of muscle tissue not selected in first zoom stage. **Blue**: larger region of muscle selected in first zoom stage but not second.
b. Kidney medulla not selected in first zoom stage.
c. Red: fibrous tissue surrounding the kidney not selected in second zoom stage.
Blue: muscle not selected in second zoom stage.

Discussion

- We designed and implemented an RL-based model that sequentially selects patches of a WSIs to localize ROIs
- We found that in the task of localizing glomeruli on kidney biopsies, the model significantly reduces the number of pixels examined while maintaining the same ROI coverage
- Two-fold contribution:
  - **Human-AI interaction**: our model has the potential to integrate into and expedite the pathologist's workflow by pre-selecting ROIs to be further analyzed by the pathologist.
  - **AI deployment and equity**: our pipeline drastically decreases the number of pixels examined, reducing the computational cost for the deployment and scaling of WSI-based models, especially in low-resource settings.
- Future directions:
  - Testing on different tasks and different tissue types
  - Incorporating and testing on more robust classifiers for a wider variety of downstream tasks after ROI selection

References

1. Burak Uzkent and Stefano Ermon. Learning when and where to zoom with deep reinforcement learning. In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, pages 12345–12354, 2020.