ComPRePS: An Automated Cloud-based Image Analysis tool to democratize AI in Digital Pathology

Sayat Mimar[a], Anindya S. Paul[a], Nicholas Lucarelli[a], Samuel Border[a], Briana Santo[b,c], Ahmed Naglah[a], Laura Barisoni[d,e], Jeffrey Hodgin[f], Avi Z Rosenberg[g], William Clapp[h], and Pinaki Sarder[a], for the Kidney Precision Medicine Project

[a]Quantitative Health Section, Division of Nephrology, Hypertension, and Renal Transplantation, Department of Medicine, University of Florida, Gainesville, FL
[b]Department of Pathology and Anatomical Sciences, University at Buffalo, Buffalo, NY
[c]Canon Stroke and Vascular Research Center, University at Buffalo, Buffalo, NY
[d]Duke University, Department of Pathology, Division of AI & Computational Pathology, Durham, NC
[e]Department of Medicine, Division of Nephrology, Duke University, Durham, NC
[f]Department of Pathology, University of Michigan, Ann Arbor, MI
[g]Department of Pathology, Johns Hopkins University, Baltimore, MD
[h]Department of Pathology, Immunology and Laboratory Medicine, University of Florida College of Medicine, Gainesville, FL

ABSTRACT

Artificial intelligence (AI) has extensive applications in a wide range of disciplines including healthcare and clinical practice. Advances in high-resolution whole-slide brightfield microscopy allow for the digitization of histologically stained tissue sections, producing gigapixel-scale whole-slide images (WSI). The significant improvement in computing and revolution of deep neural network (DNN)-based AI technologies over the last decade allow us to integrate massively parallelized computational power, cutting-edge AI algorithms, and big data storage, management, and processing. Applied to WSIs, AI has created opportunities for improved disease diagnostics and prognostics with the ultimate goal of enhancing precision medicine and resulting patient care.

The National Institutes of Health (NIH) has recognized the importance of developing standardized principles for data management and discovery for the advancement of science and proposed the Findable, Accessible, Interoperable, Reusable, (FAIR) Data Principles[1] with the goal of building a modernized biomedical data resource ecosystem to establish collaborative research communities. In line with this mission and to democratize AI-based image analysis in digital pathology, we propose ComPRePS: an end-to-end automated **Comp**utational **Re**nal **P**athology **S**uite which combines massive scalability, on-demand cloud computing, and an easy-to-use web-based user interface for data upload, storage, management, slide-level visualization, and domain expert interaction. Moreover, our platform is equipped with both in-house and collaborator developed sophisticated AI algorithms in the back-end server for image analysis to identify clinically relevant micro-anatomic functional tissue units (FTU) and to extract image features.

Keywords: AI, FAIR, Digital and Computational Pathology, Whole slide images, Functional tissue units, Image segmentation, Automatic feature extraction, AI in healthcare, Cloud platform

1. INTRODUCTION

In this work, we present ComPRePS, an open-source, cloud-based computational tool for data upload, storage, and digital histology visualization integrated with server-based AI algorithms to perform complex image analysis tasks. Our tool offers various segmentation methods to detect clinically relevant structures as well as a pipeline to

Further author information: (Send correspondence to Pinaki Sarder, PhD)
Pinaki Sarder, PhD: E-mail: pinaki.sarder@medicine.ufl.edu

extract features of clinical import. In addition, our platform has the goal of reconciling AI-based image analysis and digital pathology for clinical application. Our publicly available tool and software create the opportunity for research communities to contribute to more reliable, accurate, and explainable AI-based image analysis methods.

Digital and computational pathology is becoming increasingly popular and being adopted by institutions globally[2,3] with the promise of improving diagnostics and prognostics across organ systems. Requirements for successful implementation include addressing data management and storage of many large-scale digitized slides. In addition, high computational demand for extracting and analyzing relevant information from each WSI remains a significant bottleneck for wider adoption. With ComPRePS, we developed a platform with an intuitive browser-based user interface (UI) to improve a pathologist's interaction with WSIs and apply AI to augment clinical decision making. A few noteworthy features in ComPRePS include: (1) enabling users to upload image data with linked metadata while maintaining the highest standard of security and confidentiality, and (2) managing user data in an AI-ready format by maintaining inter-operability between popular digital pathology management such as xml, json, csv, xlsx etc. for downstream applications. In this paper, by demonstrating the segmentation task of FTUs from kidney biopsy WSIs using ComPRePS, we will describe the end-to-end pipeline to automate a time-consuming process that to date has been performed manually by an expert. We hope that ComPRePS will be the catalyst at shifting the paradigm in the research community by creating a collaborative platform where data, AI models, codes, and software are public and accessible to any user globally regardless of skill level. Our user-friendly platform enables end users including clinicians and pathologists to perform analyses on their own very large images without any AI/software training,[4] allowing all users to contribute to ever more powerful AI models to solve some of the most difficult healthcare challenges.

2. METHODS

2.1 ComPRePS User-Interface (UI)

HistomicsUI is the front-end of our tool where WSIs are organized, managed using Girder,[5] a free data management platform with visualization using Digital Slide Archive (DSA).[6] The user interface (UI) is accessible via the University of Florida Athena instance[7] of DSA (credentials: *username=public*, *password=public*) from any browser, no desktop installation required. Once the WSI is opened in the HistomicsUI, the user can interact with the digital slide as in any stand-alone WSI viewer with the ability to pan and zoom in order to view structures of interest in appropriate detail. In addition, the user can display color-coded annotation layers overlaid on the WSI itself.

2.2 ComPRePS AI pipeline for FTU segmentation

Our browser-accessible, cloud-based, user-friendly tool enables clinicians, pathologists, biologists and computational scientists to independently perform analyses with the goal of reducing their segmentation burden. To achieve this goal, our custom segmentation tool is based off of Facebook AI Research's Detectron2 library to perform complex panoptic segmentation tasks of 6 kidney-specific WSI FTUs.[8] Our segmentation pipeline can be used for training new models as well as generating predicted annotations with new user uploaded data.

For model training, we leverage the workflow of the Detectron2 library[9] and build our customized pipeline for FTU segmentation and classification, which implements convenient functions, methods and attributes by making use of a panoptic feature pyramidal architecture from input image patches. Fixed-size training patches are randomly chopped and extracted from training samples and fed into the training process. Our training data consisted of 190 WSIs with H&E, PAS, trichrome, Jones' Methenamine Silver staining, including examples of diabetic nephropathy (n=53), lupus nephritis (n=39), and transplant surveillance needle core biopsies (11), as well as 58 large sub-WSIs manually cropped from healthy portions of 33 reference kidneys, 23 small sub-WSIs of 5 H&E needle core biopsies, 2 small sub-WSIs from a silver-stained biopsy, and 4 sub-WSIs from a trichrome biopsy. Each WSI used for training had 6 associated annotation layers performed manually by domain experts: non-globally sclerotic and globally sclerotic glomeruli, cortical and medullary interstitium, tubules, arteries/arterioles.[10] Our tool also has 2 training pipelines to segment peritubular capillaries (PTC) and interstitial fibrosis and tubular atrophy (IFTA).

After the training is completed and validated on a hold-out dataset not participated in training, the trained model becomes available in ComPRePS model zoo to be applied on a prediction workflow for segmenting FTUs

on new user-uploaded WSIs. The model prediction workflow is also built by abstracting the Detectron2 library functions and methods which follow first by cropping the large input WSI into small sized patches, generating segmentation masks for each patch and finally stitching all mask patches together on the WSI level. The resulting segmentation masks are saved as JSON annotations which can be seamlessly visualized in the ComPRePS UI overlaying on the WSI.

Each segmentation pipelines are deployed in ComPRePS as a docker image including codes and dependencies by following best software engineering practices so our developed AI pipelines can be inter-operable on different servers and deployed in a quick fashion. Our user-friendly UI enables the user to select various hyper-parameters for AI tuning in order to customize training and prediction workflows according to users choice. All the pre-trained models for various FTU segmentation are saved in our default open-sourced model zoo which are ready for our users to consume directly and apply in their data. Each generated segmentation masks are displayed automatically as different color coded layers for easy visualization and quality evaluation directly on ComPRePS UI, no additional tool or application is needed. In addition, ComPRePS UI allows users to edit FTU mask boundaries as needed which are automatically saved in the JSON files. In our knowledge, currently there is no digital pathology tool available in the market which performs end-to-end (E2E) computational AI based image analysis that performs the following: First, enabling a novice user such as a bench scientist, clinician or nephropathologist to do the entire E2E slide visualization, FTU segmentation and mask generation, quality evaluation and editing AI generated masks according to their liking in a single software platform without the need of any additional third-party add-on/applications. Second, reinforcing the mission of democratizing healthcare in all over the world by harnessing the transformative power of AI in patient care with open-sourced software/code and pre-trained models for community access and feedback.

2.3 ComPRePS Pathomic Feature Extraction from Segmented FTUs

Following segmentation, the feature extraction pipeline extracts pathologist-informed morphometric features from segmented FTUs (non-globally sclerotic and globally sclerotic glomeruli, tubules and arteries/arterioles) using image analysis techniques. Extracted features are summarized as follows and can be extracted into a downloadable spreadsheet format for: **Glomeruli**: Glomerular area, mesangial area, computed mesangial fraction and bounding box coordinates for each glomerulus. **Tubules**: Average tubular basement membrane thickness, average cell thickness, luminal fraction and bounding box coordinates. **Arteries/arterioles**: Arterial area and bounding box coordinates for each segmented artery/arteriole. Each FTU feature sets have been chosen evaluating their clinical significance by consulting with several nephro-pathologists. We are currently developing an additional morphometric feature extraction pipeline that segments 3 sub-compartments (nuclei, PAS+/eosinophilic areas, and luminal space) within each FTU and extracts sub-compartment specific features.

3. RESULTS

The current list of image analysis and AI pipelines available in ComPRePS are shown in Fig.1 top panel. After selecting the desired pipeline, the user can launch either a training or prediction job by first populating the tab with hyper-parameters of interest. As an example, for the multi-compartment segmentation pipeline which segments multiple FTUs in a single pipeline: (1) user selects the image, image folder and pre-trained model and submits the job (Fig.1 lower panel), (2) upon job completion, 6 annotation layers corresponding to cortical and medullary interstitium, non-globally and globally sclerotic glomeruli, tubules, and arteries/arterioles are added under the *annotation tab* automatically with different colors and can be visualized by toggling the eye icon next to each layer (Fig.1 lower panel). The segmentation output from our pipeline produces accurate and precise segmentation results of kidney FTUs (Fig.2 panels **a-h**) which are also verified by various nephropathologists from various institutions. ComPRePS UI enables users to perform quality evaluations of the generated annotations and correct them if necessary by using various editing tools.

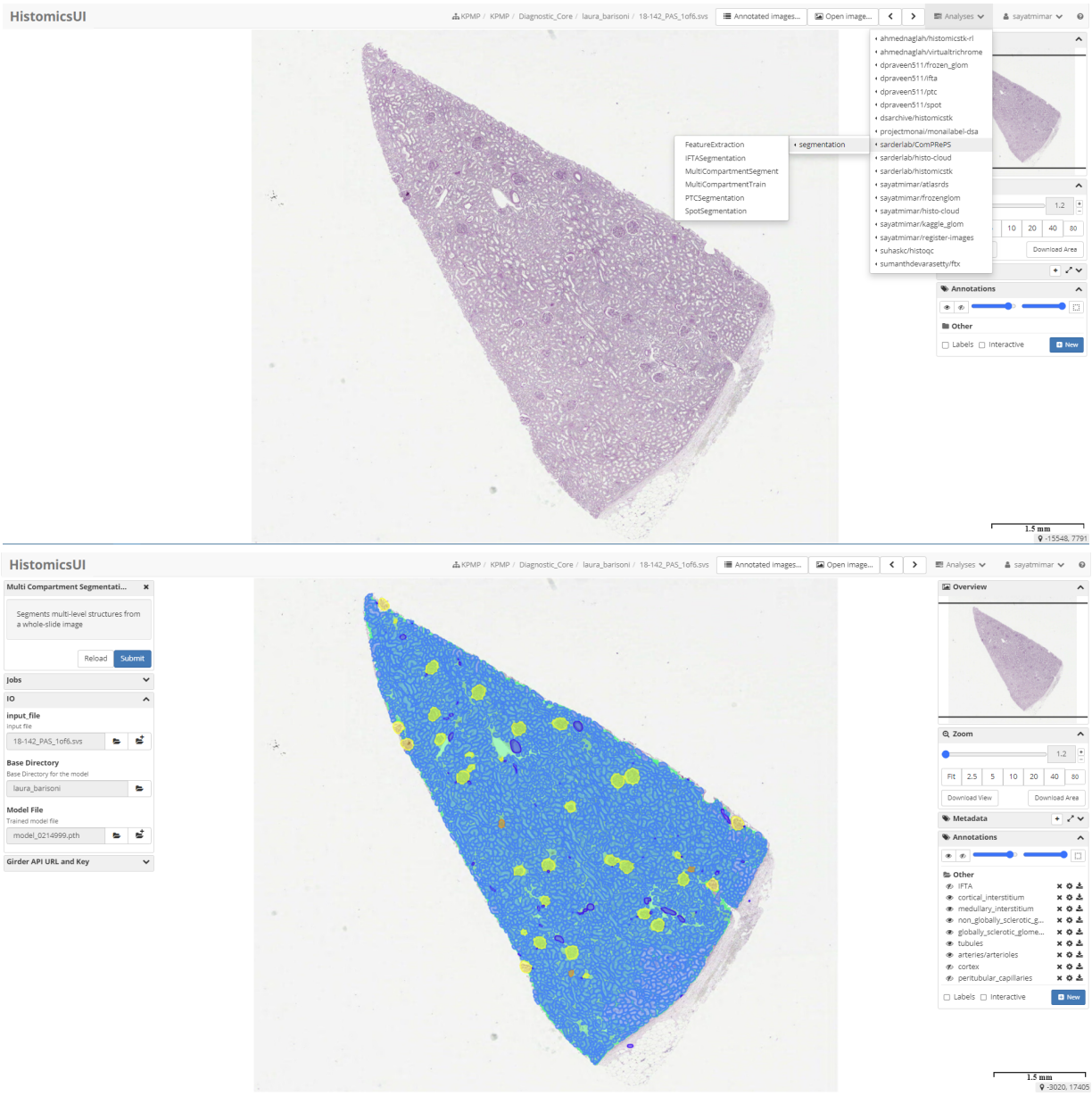
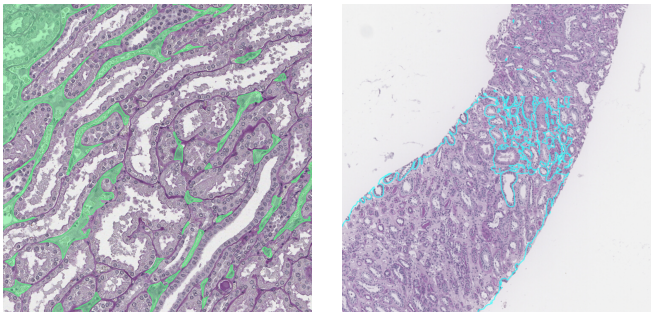
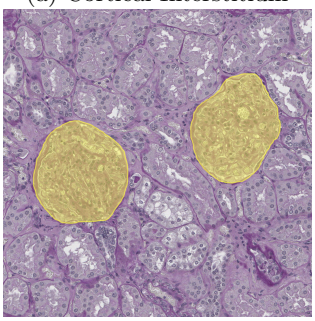


Figure 1: HistomicsUI and ComPRePS with all publicly available plugins. Multi compartment segmentation plugin shown on the top panel. The job can be run after selecting the image file, image directory and the pre-trained model. Once the job is run, annotations are added under the *annotations* tab and can be visualized by turning on/off the eye icon as shown in the lower panel.

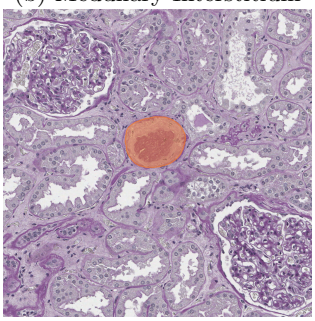


(a) Cortical Interstitium

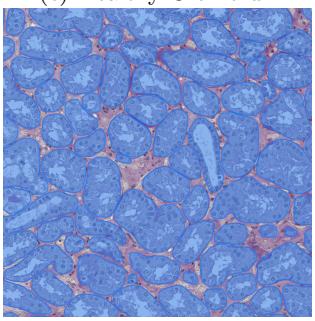
(b) Medullary Interstitium



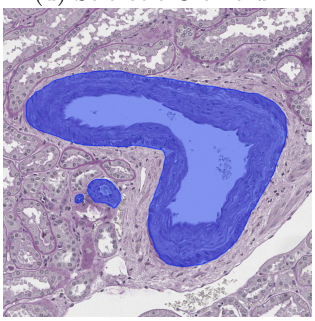
(c) Healthy Glomeruli



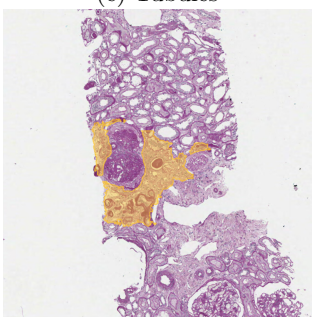
(d) Sclerotic Glomeruli



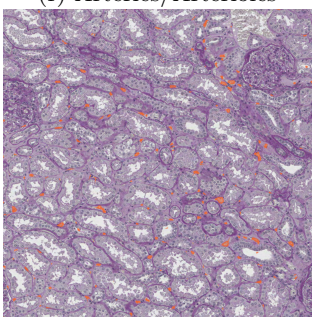
(e) Tubules



(f) Arteries/Arterioles



(g) IFTA



(h) Peritubular capillaries

Figure 2: Results of the Multi compartment segmentation model is shown in panels a-f. In panels g and h, IFTA and Peritubular capillaries segmentations are illustrated respectively. Once the job is run, annotations are added under the *annotations* tab and can be visualized by turning on/off in Histomics UI.

3.1 Real World Usability Study

As of the time of writing this manuscript, ComPRePS UI and associated AI pipelines such as FTU segmentation and feature extraction have been in active use by pathologists from multiple institutions participating in large National Institute of Health (NIH) consortia; the Kidney Precision Medicine Project (KPMP),[11] the Human BioMolecular Atlas Program (HuBMAP),[12] to visualize their scanned and uploaded WSIs and to perform quality evaluation and edit AI generated annotations. A formal large-scale usability study is planned for the near future. We invite anyone who would like to participate in the usability study to contact the corresponding author.

4. CONCLUSION

In this work, we present our automated cloud tool, ComPRePS, to perform computational image analysis tasks harnessing the life-altering power of AI to support our mission of democratizing AI in digital pathology including making the technology accessible to clinicians and pathologists so AI based findings can be augmented with human decision making during the E2E life cycle of patient care. In this article, we have reviewed the development and deployment steps of various AI pipelines in ComPRePS to perform automated model training and segmentation on WSIs. The AI pipelines integrated with our user-friendly UI simplifies the utilization process keeping end-users front and center where they can rapidly segment FTUs, generate insights and visualize results without any coding or software development experience. We open-sourced ComPRePS with all our codes and deployed docker images in GitHub and Docker Hub that is scalable to multiple users world-wide and capable to provide helpful service to clinicians and biologists whenever they need it in the clinical pipeline. In addition, ComPRePS allows researchers/practitioners to incorporate their own pre-trained model into the model zoo as well as enable users to further fine-tune their model using our growing database. Our online easily accessible platform, open-source pipelines, and scalable software can help expedite research activity as well as most importantly can make a difference in clinical diagnostics by bringing AI-generated insights to augment the expertise of clinicians and pathologists.

ACKNOWLEDGMENTS

The authors would like to acknowledge the support from the National Institute of Health (NIH) / National institute of Diabetes and Digestive and Kidney Diseases (NIDDK) grants R01 DK114485, R21 DK128668, U01 DK133090, R01 DK129541, 2R01 DK118431 and A Computational Image Analysis Platform (CIMAP) for the Human BioMolecular Atlas Program (HuBMAP) grant OT2 OD033753. The authors also acknowledge support from Nephcure Foundation and Haller Foundation.

REFERENCES

[1] Wilkinson, M. D., Dumontier, M., Aalbersberg, I. J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J.-W., da Silva Santos, L. B., Bourne, P. E., Bouwman, J., Brookes, A. J., Clark, T., Crosas, M., Dillo, I., Dumon, O., Edmunds, S., Evelo, C. T., Finkers, R., Gonzalez-Beltran, A., Gray, A. J., Groth, P., Goble, C., Grethe, J. S., Heringa, J., 't Hoen, P. A., Hooft, R., Kuhn, T., Kok, R., Kok, J., Lusher, S. J., Martone, M. E., Mons, A., Packer, A. L., Persson, B., Rocca-Serra, P., Roos, M., van Schaik, R., Sansone, S.-A., Schultes, E., Sengstag, T., Slater, T., Strawn, G., Swertz, M. A., Thompson, M., van der Lei, J., van Mulligen, E., Velterop, J., Waagmeester, A., Wittenburg, P., Wolstencroft, K., Zhao, J., and Mons, B., "The fair guiding principles for scientific data management and stewardship," *Scientific Data* **3** (Mar. 2016).

[2] Gurcan, M., Boucheron, L., Can, A., Madabhushi, A., Rajpoot, N., and Yener, B., "Histopathological image analysis: A review," *IEEE Reviews in Biomedical Engineering* **2**, 147–171 (2009).

[3] Amgad, M., Hodge, J., Elsebaie, M., Bodelon, C., Puvanesarajah, S., Gutman, D., Siziopikou, K., Goldstein, J., Gaudet, M., Teras, L., and Cooper, L., "A population-level computational histologic signature for invasive breast cancer prognosis," (May 2023).

[4] Lutnick, B., Manthey, D., Becker, J. U., Ginley, B., Moos, K., Zuckerman, J. E., Rodrigues, L., Gallan, A. J., Barisoni, L., Alpers, C. E., Wang, X. X., Myakala, K., Jones, B. A., Levi, M., Kopp, J. B., Yoshida, T., Zee, J., Han, S. S., Jain, S., Rosenberg, A. Z., Jen, K. Y., and Sarder, P. t. K. P. M. P., "A user-friendly tool for cloud-based whole slide image segmentation with examples from renal histopathology," *Communications Medicine* **2** (Aug. 2022).

[5] https://github.com/girder.

[6] https://github.com/DigitalSlideArchive/digital_slide_archive.

[7] ComPRePS implementation at the University of Florida HiPerGator: https://athena.rc.ufl.edu/.

[8] Ginley, B., Lucarelli, N., Zee, J., Jain, S., Han, S. S., Rodrigues, L., Wong, M., Jen, K. Y., and Sarder, P., "Automated reference kidney histomorphometry using a panoptic segmentation neural network correlates to patient demographics and creatinine," in [*Medical Imaging 2023: Digital and Computational Pathology*], Tomaszewski, J. E. and Ward, A. D., eds., SPIE (Apr. 2023).

[9] https://github.com/facebookresearch/detectron2.

[10] Lucarelli, N., Ginley, B., Zee, J., Mimar, S., Paul, A. S., Jain, S., Han, S. S., Rodrigues, L., Ozrazgat-Baslanti, T., Wong, M. L., Nadkarni, G., Clapp, W. L., Jen, K.-Y., and Sarder, P., "Correlating deep learning-based automated reference kidney histomorphometry with patient demographics and creatinine," *Kidney360* **4**, 1726–1737 (Nov. 2023).

[11] https://www.kpmp.org.

[12] https://hubmapconsortium.org.