National Institute of Environmental Health Sciences
Division of Translational Toxicology

Automation of Quality Control in Diagnostic Toxicologic Pathology Using Artificial Intelligence

Joshua Fletcher[1], David Cunefare[2], Charan Ganta[2], Priyanka Thakur[2], and Ronald Herbert[2]
[1]North Carolina Agricultural & Technical State University, Greensboro, NC [2]Cellular and Molecular Pathogenesis Branch, Division of Translational Toxicology, National Institute of Environmental Health Sciences, Research Triangle Park, NC 27709

Background

- Digital pathology, the process of digitizing histologic tissue sections mounted on glass slides, has permitted the use of Artificial Intelligence (AI) models such as Deep Learning (DL) to generate algorithms that assist pathologists in the accurate and efficient interpretation of pathologic lesions.
- Tissue features (i.e., staining, lesions/morphology, and architecture) can be quantifiable as visual data using computational analysis. Therefore, slides containing artifacts/imperfections (i.e., tissue folds, out-of-focus areas, and poor tissue quality) interfere with reliable data generation.
- Quality Control (QC) procedures to detect slides unsuitable for standard histopathological evaluation and AI-based analysis are usually done manually, which is a time and labor-intensive process.
- AI/DL could be used to develop automated procedures to support routine QC tasks in diagnostic toxicologic pathology.

Objectives

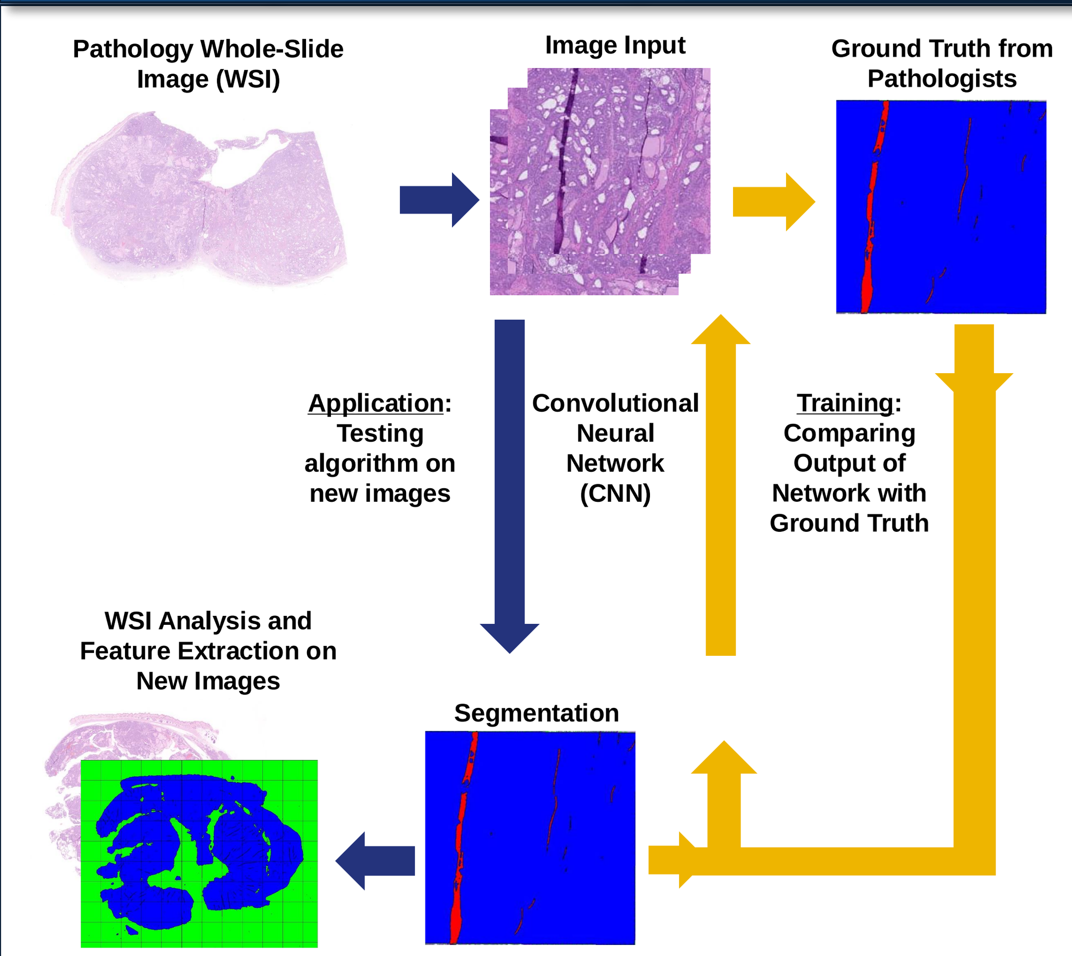
Develop Deep Learning (DL) algorithms as proof-of-concept procedures for automated Quality Control (QC) to detect:

- Scanning and tissue artifacts in whole-slide digital images using folds and out of focus artifacts as the model.
- Sub-optimal (poor quality) digitized histological tissue sections using kidney tissue sections as the model.

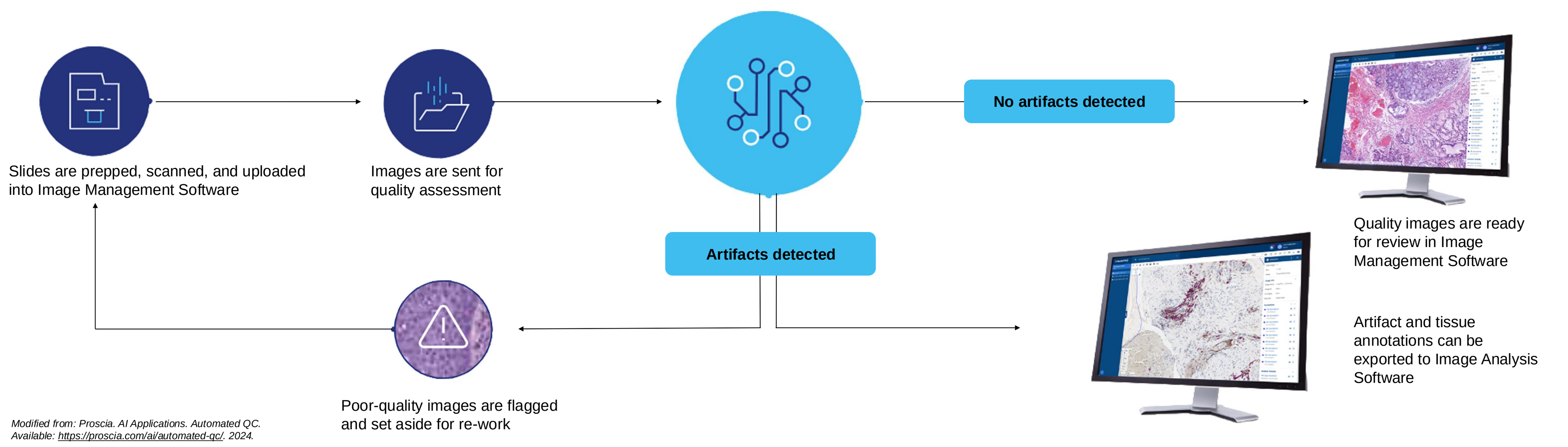
Methods

- Hematoxylin & Eosin-stained sections of 16 different rodent tissues mounted on glass slides from US National Toxicology Program studies were scanned using a digital Whole-slide Image (WSI) scanner.
- Total 44, 35, and 32 WSIs used for tissue folds, out-of-focus, and kidney groups, respectively.
- Training sets included 37, 20, and 15 slides respective to each group. Slide evaluation for training the AI applications was performed by annotating regions of interest (ROIs) on an Image Analysis Software platform.
- Two separate DeepLabv3 convolutional neural networks (CNNs) were trained to classify normal tissue, areas of artifact, and background for the tissue folds and out-of-focus groups.
- DeepLabv3 CNN was trained to classify cortex, outer medulla (outer and inner stripe), inner medulla, and background for the kidney anatomical delineation group.
- Validation was done on 10 images (5 positive controls and 5 negative controls) for each group with no overlap from the training sets. The manual grader annotated ROIs on the images, and the AI-based algorithm was run on the same ROIs.
- Calculated agreement in the validation set between manual grader vs AI.

Algorithm Training Workflow



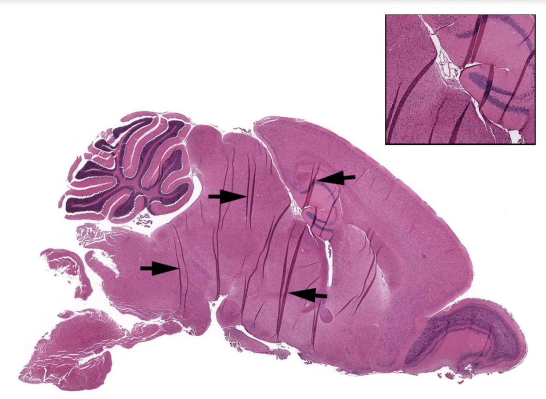
Quality Control Workflow



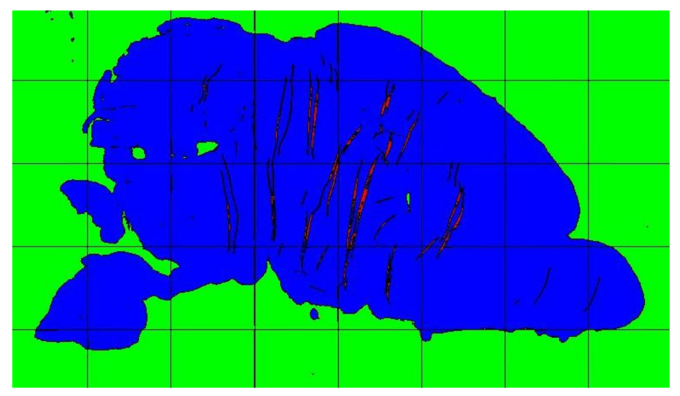
Modified from: Proscia. AI Applications. Automated QC. Available: https://proscia.com/ai/automated-qc/. 2024.

Results

H&E Section | AI Generated Annotations | Correlation with Manual Grader

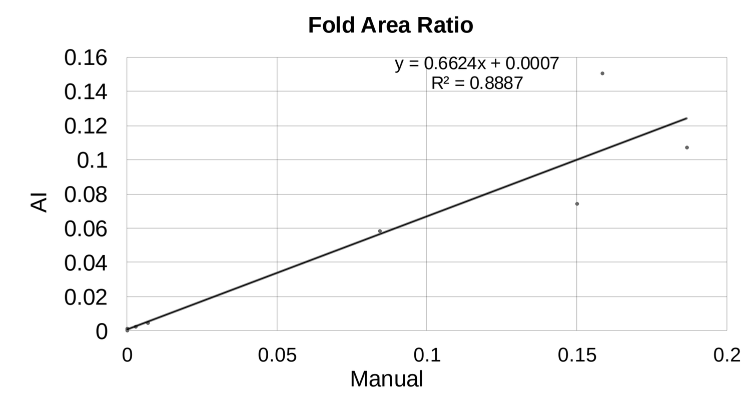


H&E-stained image of a brain section. Arrows indicate tissue fold artifacts.

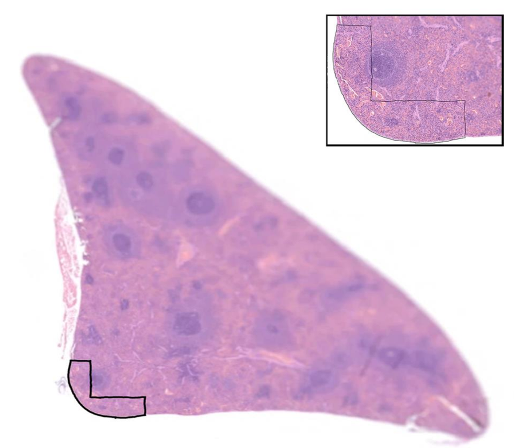


Fold detection algorithm on brain section.

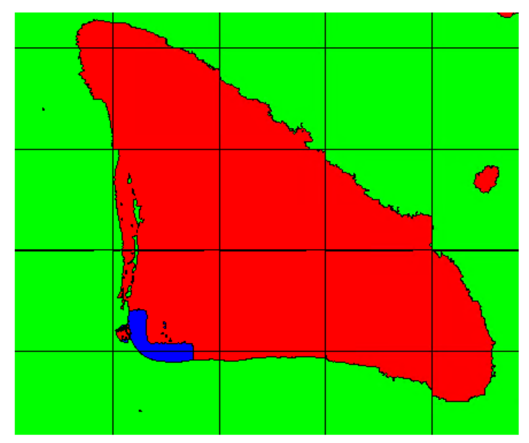
Normal Tissue, Artifact (Fold), Background



Calculated the area ratio of folds to entire tissue between manual grader and AI-based detection. The coefficient of determination (R^2) indicates correlation between detection methods.

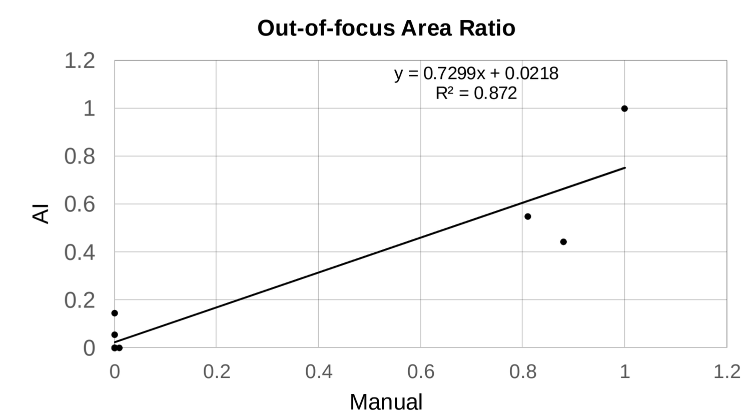


H&E-stained image of a spleen section showing out-of-focus areas with the focused area on the bottom left (inset).

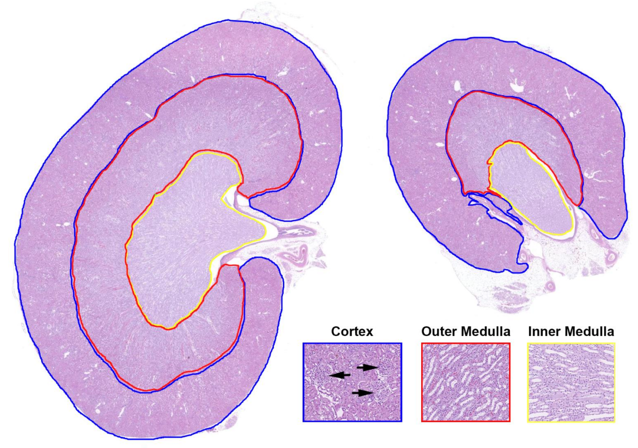


Out-of-focus detection algorithm on spleen section.

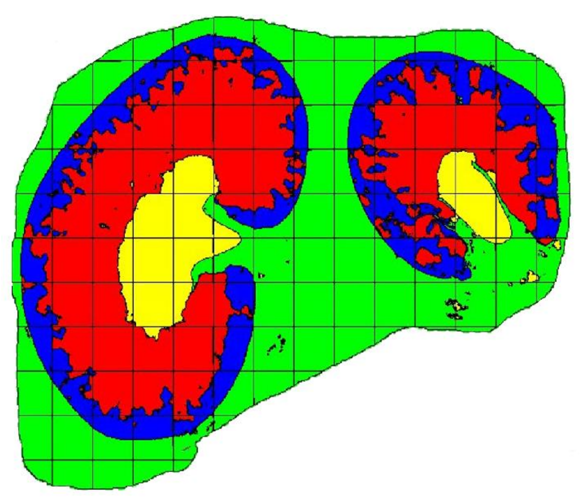
Normal Tissue, Artifact (Out-of-focus), Background

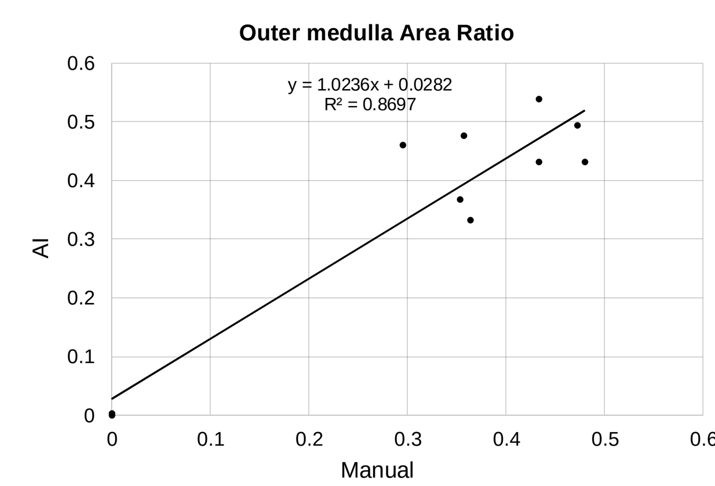


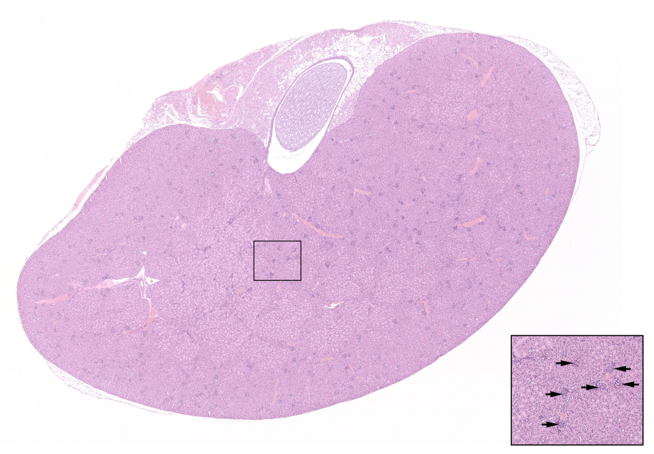
Calculated the area ratio of out-of-focus regions to entire tissue between manual grader and AI-based detection. The coefficient of determination (R^2) indicates correlation between detection methods.



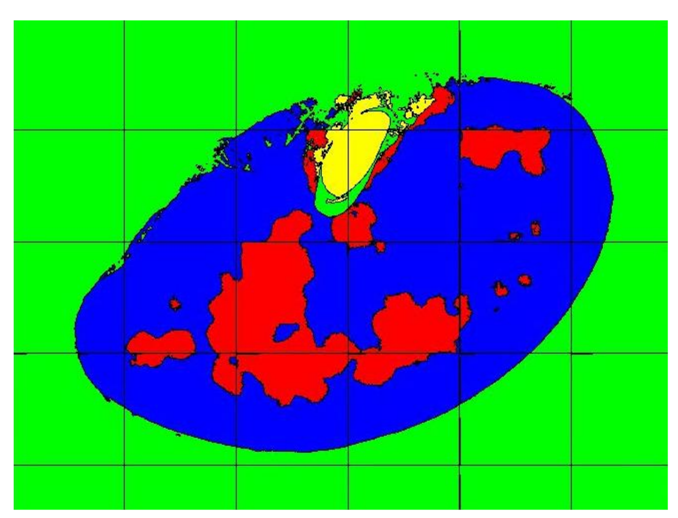
H&E-stained image of a longitudinal (left) and cross-section (right) of kidney. These are good quality sections with all anatomic regions present. See glomeruli (arrows) in cortex region.





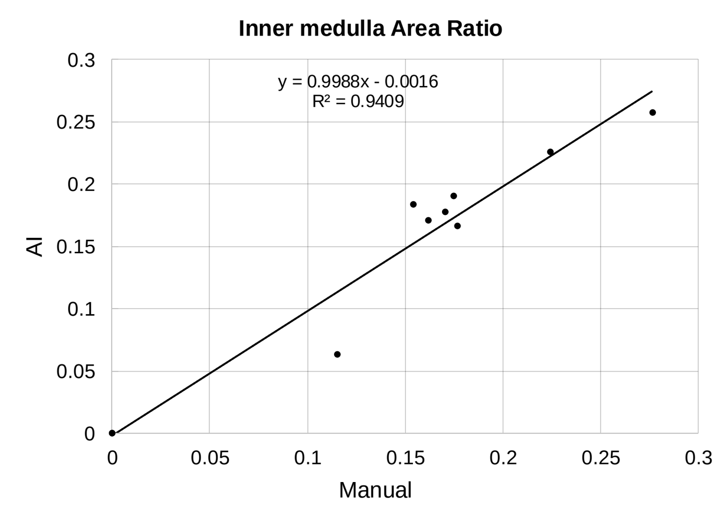


H&E-stained image of a longitudinal section of kidney. This is a sub-optimal tissue section without all the anatomical regions of the kidney clearly delineated. This section contains primarily cortex.



Anatomical region detection algorithm on kidney sections.

Cortex, Outer medulla, Inner medulla, Background



Calculated the area ratio of outer/inner medulla regions to entire tissue between manual grader and AI-based detection. The coefficient of determination (R^2) indicates correlation between detection methods.

Discussion

- The coefficient of determination (R^2) for tissue folds and out-of-focus groups are 88.9% and 87.2%, respectively, indicating a high degree of correlation between manual grading and AI-based detection.
- For kidney anatomical delineation, the outer medulla and inner medulla show R^2 values of 87% and 94.1%, respectively, suggesting the algorithm can detect the anatomical structures of the kidney with a high degree of correlation between manual annotations and AI-based segmentation.
- Increasing images in the training sets will improve AI-based detection that reduces the over/under-detection of artifacts (i.e., false positives and false negatives).

Conclusion

- This study demonstrates a proof-of-concept procedure for artifact detection in tissue folds, out-of-focus areas, and anatomical region sectioning errors, and can be implemented in the digital pathology workflow.
- These procedures will be used to establish more rigorous algorithms that can detect a variety of artifacts (i.e., mounting artifacts, knife cuts, chatter, and cracks).
- Test data from the DL algorithms can flag errors before reaching pathologists, providing automated support to QC tasks and ultimately improving diagnostic accuracy and efficiency.
- Test data can provide routine feedback to histotechnicians and reinforce high standards for QC as diagnostic toxicologic pathology is transitioning to digitalization (Garbage in, Garbage out).
- Using kidney as a model for detecting tissue anatomical artifacts, similar algorithms can be developed to detect tissue sectioning anatomical artifacts in other tissues.

References

Kim I, Kang K, Song Y, Kim T.-J. Application of Artificial Intelligence in Pathology: Trends and Challenges. Diagnostics. 2022 Nov, 12112794.

Proscia. AI Applications. Automated QC. Available: https://proscia.com/ai/automated-qc/. 2024.

Rashidi HH, Tran NK, Betts EV, Howell LP, Green R. Artificial Intelligence and Machine Learning in Pathology: The Present Landscape of Supervised Methods. Acad Pathol. 2019 Sep 3;6:2374289519873088.

Xu J, Luo X, Wang G, Gilmore H, Madabhushi A. A Deep Convolutional Neural Network for segmenting and classifying epithelial and stromal regions in histopathological images. Neurocomputing. 191: 214-233.

Wang S, Yang D, Rong R, Zhan X, Xiao G. Pathology Image Analysis Using Segmentation Deep Learning Algorithms. Am J Pathol. 2019 Sep; Am J Pathol. 2019 Sep; 189: 1686-1698.

Zuraw A, Staup M, Klopfleisch, Aeffner F, Brown D, Westerling-Bui, Rudmann D. Developing a Qualification and Verification Strategy for Digital Tissue Image Analysis in Toxicologic Pathology. Toxicol Pathol. 2021 Jun;49(4):773-783.

Zuraw A. 2024. Digital Pathology 101. 1st ed. Delaware

Acknowledgments

The authors would like to thank Eli Ney and Ashley Paragone from the CMPB Image Analysis and Artificial Intelligence Cores for generating the whole slide digital image scans for this study, and Quashana Brown and Renee Fisher from Histology and Necropsy Core for their support in preparing the H&E-stained histology slides.