www.nephropathol.com DOI: 10.34172/jnp.2021.xx J Nephropathol. 2021;10(x):exx

Journal of Nephropathology

A deep-learning approach at automated detection of electron-dense immune deposits in medical renal biopsies

Alaa Alsadi[1*], Nasma K. Majeed[2], Dereen M. Saeed[2], Yash Dharmamer[2], Manmeet B. Singh[2], Tushar N. Patel[2]

[1]Department of Pathology, University of Wisconsin, Madison, Wisconsin, USA
[2]Department of Pathology University of Illinois, Chicago, Illinois, USA

ARTICLE INFO

Article type:
Original Article

Article history:
Received: 14 October 2020
Accepted: 16 November 2020
Published online: 28 November 2020

Keywords:
Kidney
Deep learning
Biopsy
Machine learning
Electron microscopy

ABSTRACT

Introduction: Identification of electron-dense immune deposits in electron microscopy (EM) images is integral to the diagnosis of medical renal disease. Deep learning has the potential to augment this process, especially in areas with limited resources.
Objectives: Our study explores the feasibility of applying deep learning to detect electron dense immune deposits in electron microscopy images from medical renal biopsies.
Patients and Methods: EM images (N=900) from native and transplant kidney biopsies were processed into 4530 tiles (512 x 512 pixels). These tiles were reviewed and classified into one of three categories: deposits absent, deposits present, and indeterminate. This classification resulted in 1255 images with consensus agreement for deposits present and deposits absent. These 1255 images were then used to train a machine learning model, using 1006 images for training, and 249 images for testing.
Results: The overall accuracy on the test data was a competitive 78%, and the F1 scores for deposits absent and present was 0.76 and 0.79, respectively.
Conclusion: This study demonstrated the feasibility of creating and applying a machine learning model that performs competitively in identifying electron dense deposits in EM images.

Original Article

Implication for health policy/practice/research/medical education:
This study demonstrates novel application of deep learning towards analysis of EM images in the diagnosis of renal disease. It also demonstrates feasibility of introducing artificial intelligence/machine learning concepts into pathology residency training programs, especially those with low resources.

Please cite this paper as: Alsadi A, Majeed NK, Saeed DM, Dharmamer Y, Singh MB, Patel TN. A deep-learning approach at automated detection of electron-dense immune deposits in medical renal biopsies. J Nephropathol. 2021;10(x):exx. DOI: 10.34172/jnp.2021.xx.

Introduction

The diagnosis of medical renal disease involves careful consideration and integration of clinical data and morphological features. The renal pathologist routinely utilizes light microscopy, immunofluorescence studies, and electron microscopy (EM) to render a diagnosis. Electron microscopy is most commonly used in renal pathology for the evaluation of glomerulopathies, but also in a variety of other neoplastic and non-neoplastic diseases in other organ systems (1) (Table 1). In particular for renal pathology, EM facilitates ultrastructural examination of the glomerular, tubular, and vascular compartments to a degree not possible using other modalities. Diagnostic advantages of EM include: 1) Detecting glomerular injury, 2) utility in small or suboptimal tissue samples, and 3) increased resolution as compared to other methods, which enable pathologists to identify, further localize, and/or characterize electron-dense deposits along glomerular and/or tubular basement membranes (2). In a study of 88 cases including 79 native kidney biopsies and 8 allograft kidney biopsies, EM was found to have an important diagnostic role in 75% of cases, and essential or necessary in 25% of the cases (3). In a series of 115 native kidney biopsies, EM was found to be crucial for the diagnosis or have an important contribution in 12% and 20% of the cases, respectively (4). Although recent advances in ancillary techniques such as molecular testing will undoubtedly enhance our role in diagnosing

*Corresponding author: Alaa Alsadi, Email: alsadi@wisc.edu

Table 1. List of entities for which electron microscopy aids in diagnosis

Non-Neoplastic	Neoplastic
Glomerulopathies	Mesothelioma vs. Adenocarcinoma
Microbial diseases	Soft tissue tumors
Cilia abnormalities	Gastrointestinal tumor (GIST)
Microvillous inclusion diseases	Clear cell ependymoma
Lysosomal storage diseases	Dendritic reticulum cell sarcoma
Bullous skin disorders	True oncocytoma
CADASIL	Granular renal epithelial tumors
Peripheral neuropathies	Unknown primary
Striated muscle diseases	

renal diseases, EM continues to be integral to pathology practice.

Despite routine and wide usage, EM has limitations, most notably in the assessment of glomerular diseases. Certain glomerular diseases, including pauci-immune and anti-glomerular basement membrane, do not have specific ultrastructural findings. Typical procurement of tissue for renal biopsy typically involves collecting 1mm cubes from the ends of needle core biopsies (5). As a result, affected areas for diseases that are focal or not otherwise diffusely penetrative may not be adequately sampled. Finally, EM does not allow for distinguishing the biochemical specificity of ultrastructural lesions (2). This limitation poses a challenge not only when attempting to distinguish the nature of immune-complexes visualized, but also between artifactual or benign entities that look similar to pathologic lesions. Our study seeks to address this latter shortcoming by leveraging deep learning to distinguish between immune-complex electron-dense deposits and other features, such hyaline, mesangial interposition, or artifact.

Objectives

In this study, we explore a deep learning approach to facilitating this challenging and time-consuming task. Electron microscopic examination requires examination of static, high-resolution digital images, and is therefore well-suited for image recognition applications enabled by deep learning. Paucity of deep learning applications in renal pathology (especially compared with radiology and other areas of pathology) may be attributed to the complexity of implementing and the difficulty in acquiring necessary resources to develop a clinically relevant deep learning platform (6). However, recent studies suggest that this is changing and increasing attention is being given to machine learning applications in nephrology (7).

We demonstrate a potential application of supervised deep learning algorithms for renal pathology practice, using mostly publicly available hardware and software.

Patients and Methods

Platform

We used an open-source neural network library (Keras) running on top of TensorFlow (machine learning library). The hardware utilized was NVIDIA GTX 1080ti Graphical Processing Unit (GPU) with 11GB of RAM on an Intel Core i7 @ 3.7GHz CPU system with 32GB of RAM running Linux Ubuntu 16.04LTS.

Imaging data; pre-processing and labeling

We identified approximately 30 000 tagged image file format (TIFF) EM images from both native and allograft renal biopsies collected between 2003 and 2018 at the University of Illinois Hospital and Health Sciences System, Chicago, IL (Figure 1). From these 30 000 images as a starting point, and to ensure our images contain the relevant morphology needed for electron dense deposit identification, we performed natural language search in our lab information system (Cerner PathNet) to isolate cases of lupus nephritis, IgA nephropathy, and membranous nephropathy. This search resulted in isolation of 900 high-resolution images (3120 x 2190 pixels) with confirmed immune deposits by EM on final diagnosis reports. Using Python scripting, each de-identified high-resolution image was broken down in smaller tiles (512 × 512 pixels each). This process resulted in total of 4530 tiles (Figure 2A). The practice of tiling large (e.g. most medical) imaging data is common in deep learning practices due to its benefits of preventing loss of image spatial information as well as providing easier control on the parameters of the deep learning process while reducing the required computational power and memory needed for the task (8). These tiled images were subsequently converted to ".jpg" format to facilitate OpenCV software manipulation operations on the data (e.g., tiling). The tiles were subsequently uploaded to a platform that allowed facile classification of tiled images (Figure 2B).

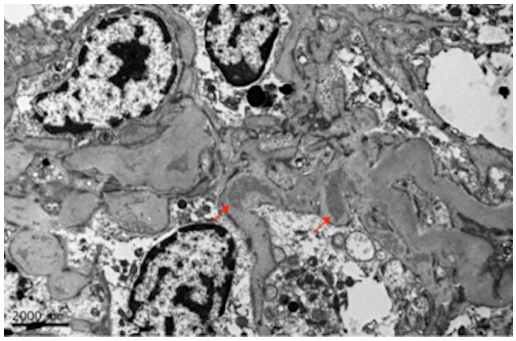

Figure 1. Sample image generated from our EM device. The image was taken from a patient with Lupus Nephritis where subendothelial deposits (arrows mark larger ones) are evident throughout. Image was generated natively as a .tif file with a resolution of 3120x2190 pixels.

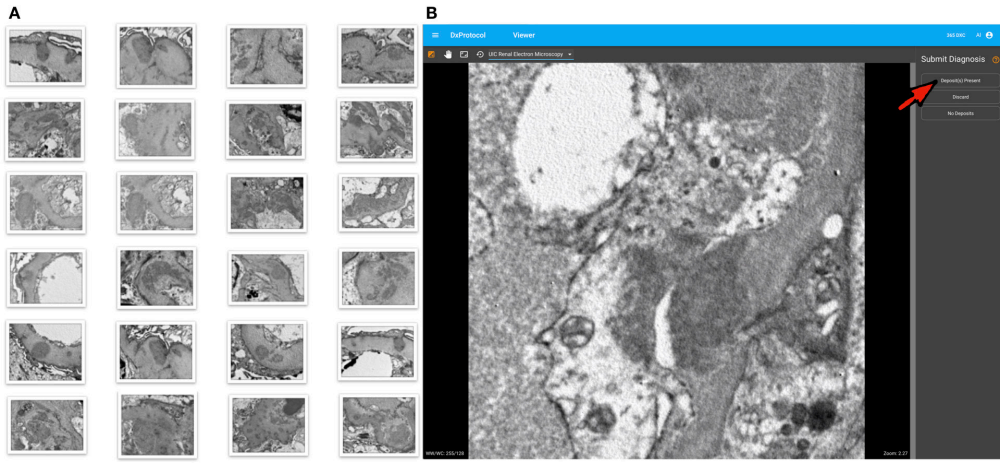

Figure 2. A. Extracted 512 × 512 pixel tiled images from high-resolution EM images of cases with known deposits. B. Process of classifying tiled images into deposit present, deposits absent, and indeterminate.

Two pathology residents with prior EM imaging experience (technicians) and one attending nephropathologist categorized EM deposits on 1640 of the total 4530 tiles into one of 3 categorized based on presence/absence of deposits: Present, Absent, and Indeterminate. Only images for which there was consensus agreement (3/3) for positive and negative were used for training and validation. The resulting training data set included 1006 images, and the validation data set included 249 images. Of the 1006 images in the training set, 512 images had deposits present, while in 494 images, deposits were absent (Figure 3).

Training

Experimental run was made using the deep neural architecture for VGG16, VGG19, InceptionResNetV2, ResNet50, NASNet and Inception-v3 networks as pre-trained in the Keras Python package (8). Utilizing these pre-trained network, we compared the effectiveness of the neural network applied at the feature extraction task, and when transferring the network (using some of the parameters) to renal pathology image classification (9). For fine-tuning these models, the setup parameters were varied to give the optimal results for this particular application. Using the above pre-trained networks and applying their layers to extract features from pathological data yielded better performance (9). Through validation it was decided that VGG19 appears to demonstrate better performance (10). After the final convolutional blocks in VGG19, we added convolutional layers that would be re-trained. These layers serve as simple attention maps that remove parts of the image that do not contribute to the classification (11). These layers apply the attention equally to all features in that patch. Instead of using classical global max pooling and average pooling layers between convolutional layers where it indicates the precise regions of interest, we used rescaled global average pooling (12). Subsequently, one fully connected layer of size 512 (followed by an output layer) was used to replace the default VGG19 fully connected layers while performing hyper parameter tuning to give increased performance (13). The optimization algorithm applied was stochastic gradient descent following the methodology whereby the step size used was small and as the learning of the model stagnated it was decreased to a very small value 0.00656. The momentum parameter of Nesterov variety was set to 0.9. Both these parameters were varied gradually to make

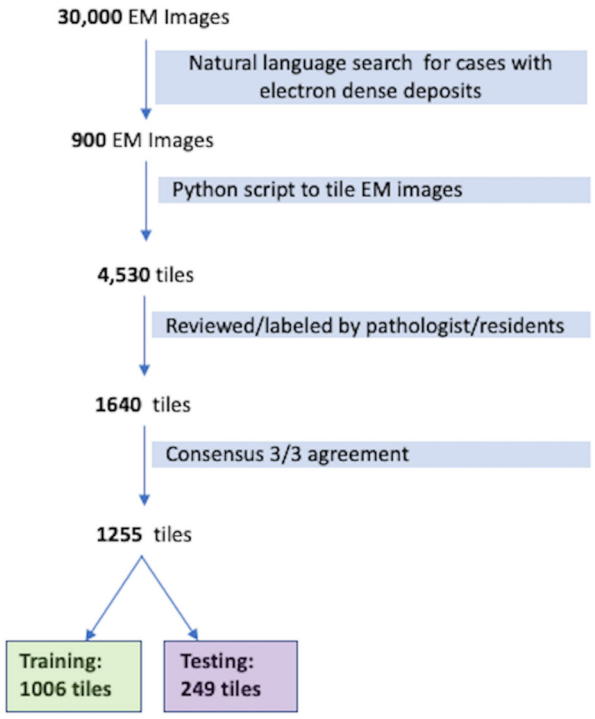

Figure 3. Overview of imaging data acquisition and labeling.

sure that network did not drastically change during the training period (13). The model was trained for a total of 15 epochs until the point that there was no appreciable change in accuracy with a batch size of five (14). To avoid computations on the CPU becoming a bottleneck, we used Tensorflow based image augmentation function that runs on GPU to generate extra training samples using various transformations such as horizontal and vertical flips, rotations, and crops (15).

Ethical issues

This investigation was in accordance with the Declaration of Helsinki. The study was performed retrospectively on EM images of kidney biopsies obtained at a single institutional center. This study was approved by the Institutional Review Board at the University of Illinois at Chicago (UIC).

Results

Testing the performance of the model was an integral part and was performed on the validation data set (249 images) on the Keras platform. Results were visualized by model performance as compared to pathologist classification (Figure 4).

When EM images contained deposits, our model had a recall (sensitivity) of 0.86 and a precision (positive predictive value) of 0.72. In other words, our model will flag 86% of the EM images deemed positive (containing deposits) by human renal pathologists. Also, for each positive flag by our model, there is a 72% probability of concordance with a renal pathologist.

When EM images contained no deposits, our model had a recall (sensitivity) of 0.69 and a precision (positive predictive value) of 0.84. In other words, our model will identify 69% of the EM images deemed negative (lacking deposits) by renal pathologists. Additionally, for each negative classification by our model, there is an 84% probability of concordance with a renal pathologist.

F1 score, a measure that combines recall and precision (harmonic mean of both), was 0.79 and 0.76 for presence of deposits and absence of deposits, respectively.

Overall accuracy was 0.78. In other words, 78% of test images were labeled correctly and same as a renal pathologist would. Area under the curve was 0.85 (Figure 5). Results are summarized in Table 2.

Discussion

Glomerular, tubulointerstitial, and vascular diseases often require ultrastructural examination for their appropriate diagnosis and classification. Several ultrastructural features are essential for appropriate interpretation of EM images: presence of organized deposits, the quality/texture of the deposits, and the relative location of the deposits. Complex medical comorbidities (e.g., transplant), inherent variability in processing workflow for EM, artifactual lesions due to processing, and gray-scale nature of the imaging make it difficult to definitively identify organized electron dense deposits from mimickers such as artifact, hyaline and mesangial interposition (16). For this reason, we sought to investigate the feasibility of a supervised deep learning approach for EM image analysis challenges.

Electron microscopic studies are well-suited for machine learning applications, since the studies are acquired in high resolution digital format. This feature is in contrast to hematoxylin & eosin (H&E) stained glass slide images where the primary product is a glass slide that secondarily undergoes digitization through scanning. For this reason, we see EM as a suitable, cost effective, application for pilot deep-learning deployment in clinical practice for renal pathologists (17).

We set to prototype a clinically relevant deep learning approach to help in detecting electron-dense immune deposits in medical renal biopsies. As a first step in this direction and to start with an achievable scope, we aimed to train a model that can distinguish between the presence or absence of organized deposits in EM images. Our primary aim was to develop an approach that is implementable

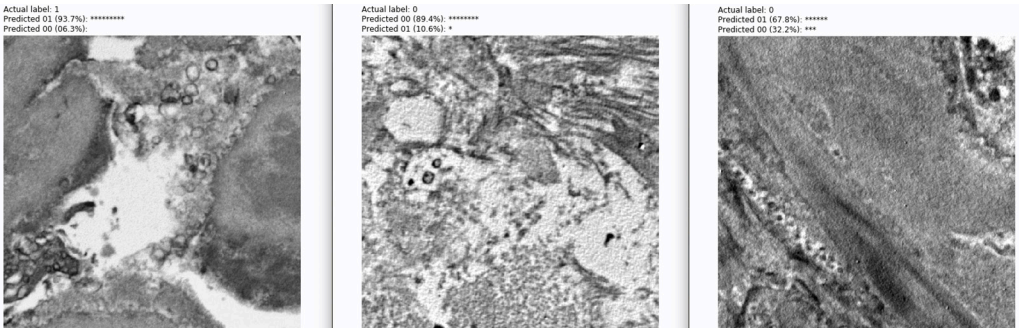

Figure 4. Sample images from model performance on testing images, that were reviewed by renal pathologist. Each image demonstrates the model's predicted label and the actual label assigned by pathologist. 0 = deposits absent; 1 = deposits present.

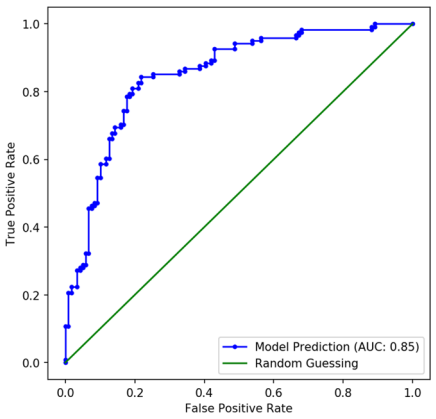

Figure 5. Model prediction area under the curve (AUC)

using accessible resources, such as commercially available deep learning frameworks and workstations. Additional insights and answers to more complex questions (eg, localization of deposits) can certainly be gleaned with more advanced resources, including additional data, computing power, and clinical expertise.

Our results were achieved with minimal resources and primarily intra-departmental staffing, including residents, graduate students, and fellows. Aside from proving the feasibility of this approach, the accessible resources utilized further enables the incorporation and application of machine learning concepts into training programs. The use of commercially available software/hardware that is approachable for the non-technical individual lowers the apprehension trainees often feel when attempting to involve themselves in machine learning projects. There is increasing support for training in machine learning to be incorporated into pathology residency programs to prepare residents for employment positions (18).

The performance of our machine learning model is competitive within the context of relatively limited data set that was assembled. Nonetheless, much can still to be done for improving the performance and generalizability through enriching the data set (e.g. including subsets of artifacts and resorbed deposits) and fine-tuning neural network parameters and additional customization of the neural network. Segmentation of electron-dense deposits within training data would further enhance the specificity of classification and feature identification. Enriching the data with extra-institutional images would increase generalizability of the model as well. Finally, including a broader set of disease entities would also contribute to robustness of model performance if a correlation exists between disease and morphologic appearance of electron dense deposits. Overall, our findings suggest that deep learning algorithms have potential for facilitating the task of identifying organized deposits in EM images, and thereby aiding the renal pathologist in their workflow.

Table 2. Summary of performance metrics

Deposits	Precision	Recall	F1-score	Image count
Absent	0.84	0.69	0.76	124
Present	0.72	0.86	0.79	116

Accuracy on test data: 0.78

While our prototype was limited to a two classes classifier, future potential applications of this deep learning approach could include a localization and/or further characterization of the deposits. An interface for such functionality could overlay bounding boxes or segmentations or contours around these deposits, an approach commonly applied to scanned H&E images (19). Further calculations to aid in deciding predominant distributions (e.g., subepithelial versus subendothelial) could also be implemented. In addition to deposit localization, models can be trained to classify the composition of the deposits (e.g., amyloid versus non-amyloid), based on high-power morphological appearance of fibrils (e.g, diameter, organization, etc.). Moreover, investigating the molecular nature of these deposits (e.g. amyloid AA vs amyloid AL, etc.) could prove to be cost and time saving; for this scenario data sets should be matched to those of molecular/proteomic studies. The challenge for these additional explorations includes gathering sufficient high-quality data sets to enable robust training and validation.

Most important future effort, however, should focus on dedicating resources to validate such a prototype in the actual clinical workflow. This could start with a pilot to execute our algorithm on clinical EM images and share results with the renal pathologist on service. Registering the feedback from practicing renal pathologists and fine tuning to their preferences could be the first step in promoting the utility of our approach and deep-learning approaches in general. Workflows for such incorporation do not have to be complex. In our institution, EM images are stored on a network drive outside of the electronic medical record and radiology picture archiving and communication system. The EM images are viewed with basic operating system image viewers. Incorporating a deep-learning algorithm in the clinical workflow could be as simple as having a technical staff member from the EM lab make a copy of EM images for a certain case, run the prediction of the trained algorithm in a standalone environment, and share the results with the renal pathologist as a second copy of the EM images for their review. Ideally, the electron medical record would contain the images and the trained algorithms, allowing the pathologist to workflow to be seamless and less prone to error.

Finally, departmental lead initiatives are of paramount

importance to set adoption of this technology to success. In a 2019 survey of the global pathology community, around 80% of the respondents indicated their belief that artificial intelligence technology would be incorporated into pathology/laboratory practice within ten years (20). Given the exponential academic and translational interest in applying machine learning to improve patient care, training programs are recognizing the importance of incorporating these concepts into training pathway. Studies that include readily available and accessible resources such as ours allow trainees from a variety of background to develop familiarity with the benefits and limitations of this emerging technology.

Conclusion

Our results provide support for the potential use of deep learning approach as a tool to identify electron-dense immune deposits in medical kidney disease. Factors promoting adoption of deep learning methods into renal pathologist practice include: EM images natively digital acquisition, increasing availability of deep learning platforms (especially from commercial sector), and decreasing costs of hardware (eg, GPU) and cloud-based solutions. Integration of machine learning models should ideally be done with the electronic medical record; however, the short-term roadmap likely will employ these models outside of the electronic medical record. Familiarity with machine learning methods and working within teams that leverage machine learning is increasingly being recognized as skills of future physician leaders.

Limitations of the study

The study was limited primarily by data set constraints. This included both the total volume of images collected, as well as the variety of disease entities represented by the images. Furthermore, images from multiple institutions would contribute the generalizability of the model.

Authors' contribution

All authors made substantial contributions to the acquisition (AA and MBS), processing (YD, AA), and classification (NKM, DMS and TNP) of the data. All authors participated in the drafting or revision of work, and gave their final approval to the submitted version.

Conflicts of interest

The authors declare no competing interests.

Ethical considerations

Ethical issues (including plagiarism, data fabrication, double publication) have been completely observed by the authors.

Funding/Support

The authors would like to thank the UIC department of pathology for the funding of this project.

References

1. Erlandson RA. Role of Electron Microscopy in Modern Diagnostic Surgical Pathology. Weidner N, Cote RJ, Suster S, Weiss LM, eds. Mod Surg Pathol. 2009:71-84. doi: 10.1016/B978-1-4160-3966-2.00005-9.
2. Pavlisko EN, Howell DN. The continued vital role of electron microscopy in the diagnosis of renal disease/dysfunction. Ultrastruct Pathol. 2013;37:1-8. doi: 10.3109/01913123.2012.670025.
3. Darouich S, Goucha RL, Jaafoura MH, Moussa FB, Zekri S, Maiz HB. Value of Electron Microscopy in the Diagnosis of Glomerular Diseases. Ultrastruct Pathol. 2010;34:49-61. doi: 10.3109/01913120903506637.
4. Pearson JM, McWilliam LJ, Coyne JD, Curry A. Value of electron microscopy in diagnosis of renal disease. J Clin Pathol. 1994;47:126-128. doi: 10.1136/jcp.47.2.126.
5. Walker PD, Cavallo T, Bonsib SM, Ad Hoc Committee on Renal Biopsy Guidelines of the Renal Pathology Society. Practice guidelines for the renal biopsy. Mod Pathol Off J U S Can Acad Pathol Inc. 2004;17:1555-1563. doi: 10.1038/modpathol.3800239.
6. Chartrand G, Cheng PM, Vorontsov E, Drozdzal M, Turcotte S, J Pal C, et al. Deep Learning: A Primer for Radiologists. Radiogr Rev Publ Radiol Soc N Am Inc. 2017;37:2113-31. doi: 10.1148/rg.2017170077.
7. Lemley KV. Machine Learning Comes to Nephrology. J Am Soc Nephrol JASN. 2019;30:1780-1781. doi: 10.1681/ASN.2019070664.
8. Bandi P, Geessink O, Manson Q, van Dijk M, Balkenhol M, Hermsen M, et al. From Detection of Individual Metastases to Classification of Lymph Node Status at the Patient Level: The CAMELYON17 Challenge. IEEE Trans Med Imaging. 2019;38(2):550-60. doi: 10.1109/TMI.2018.2867350.
9. Hoo-Chang S, Roth HR, Gao M, Le Lu Ziyue, Nogues I, Yao J, et al. Deep Convolutional Neural Networks for Computer-Aided Detection: CNN Architectures, Dataset Characteristics and Transfer Learning. IEEE Trans Med Imaging. 2016;35:1285-1298. doi: 10.1109/TMI.2016.2528162.
10. Simonyan K, Zisserman A. Very Deep Convolutional Networks for Large-Scale Image Recognition. ArXiv14091556 Cs. Published online September 4, 2014. http://arxiv.org/abs/1409.1556. Accessed March 11, 2019.
11. Attention in Neural Networks and How to Use It. http://akosiorek.github.io/ml/2017/10/14/visual-attention.html. Accessed September 28, 2019.
12. Qiu S. Global Weighted Average Pooling Bridges Pixel-level Localization and Image-level Classification. ArXiv180908264 Cs. Published online September 21, 2018. http://arxiv.org/abs/1809.08264. Accessed September 28, 2019.
13. Girshick R, Donahue J, Darrell T, Malik J. Region-Based Convolutional Networks for Accurate Object Detection

and Segmentation. IEEE Trans Pattern Anal Mach Intell. 2016;38:142-158. doi: 10.1109/TPAMI.2015.2437384.
14. Keras-Team/Keras. Keras; 2019. https://github.com/keras-team/keras. Accessed September 28, 2019.
15. Arkhangelskiy S. Data augmentation on GPU in Tensorflow. Medium. Published November 28, 2017. Accessed September 24, 2019. https://becominghuman.ai/data-augmentation-on-gpu-in-tensorflow-13d14ecf2b19.
16. Guillermo A. Herrera, Elba A. Turbat-Herrera; Renal Diseases With Organized Deposits: An Algorithmic Approach to Classification and Clinicopathologic Diagnosis. Arch Pathol Lab Med 2010; 134:512-31.
17. Janowczyk A, Madabhushi A. Deep learning for digital pathology image analysis: A comprehensive tutorial with selected use cases. J Pathol Inform. 2016;7:29. doi: 10.4103/2153-3539.186902.
18. Green R, Hogarth MA, Prystowsky MB, Rashidi HH. The Job Market Outlook for Residency Graduates: Clear Weather Ahead for the Butterflies? Arch Pathol Lab Med. 2018;142:435-438. doi: 10.5858/arpa.2017-0540-ED.
19. Wang D, Khosla A, Gargeya R, Irshad H, Beck AH. Deep learning for identifying metastatic breast cancer. ArXiv 160605718. Published online 2016. https://arxiv.org/abs/1606.05718. Accessed August 30, 2017.
20. Sarwar S, Dent A, Faust K, et al. Physician perspectives on integration of artificial intelligence into diagnostic pathology. Npj Digit Med. 2019;2:28. doi: 10.1038/s41746-019-0106-0.