Development and Validation of a Deep-learning Model to Assist With Renal Cell Carcinoma Histopathologic Interpretation

Michael Fenstermaker, Scott A. Tomlins, Karandeep Singh, Jenna Wiens, and Todd M. Morgan

OBJECTIVE To develop and test the ability of a convolutional neural network (CNN) to accurately identify the presence of renal cell carcinoma (RCC) on histopathology specimens, as well as differentiate RCC histologic subtype and grade.

MATERIALS AND METHODS Digital hematoxylin and eosin stained biopsy images were downloaded from The Cancer Genome Atlas. A CNN model was trained on 100 um2 samples of either normal (3000 samples) or RCC (12,168 samples) tissue samples from 42 patients. RCC specimens included clear cell, chromophobe, and papillary histiotypes, as well as tissue of Fuhrman grades 1 through 4. Model testing was performed on an additional held-out cohort of benign and RCC specimens. Model performance was assessed on the basis of diagnostic accuracy, sensitivity, specificity, positive predictive value, and negative predictive value.

RESULTS The CNN model achieved an overall accuracy of 99.1% in the testing cohort for distinguishing normal parenchyma from RCC (sensitivity 100%, specificity 97.1%). Accuracy for distinguishing between clear cell, papillary, and chromophobehistiotypes was 97.5%. Accuracy for predicting Fuhrman grade was 98.4%.

CONCLUSION CNNs are able to rapidly and accurately identify the presence of RCC, distinguish RCC histologic subtypes, and identify tumor grade by analyzing histopathology specimens. UROLOGY 144: 152–157, 2020. © 2020 Elsevier Inc.

Renal cell carcinoma (RCC) outcomes, treatment options, and surveillance regimens are highly dependent upon accurate staging, including the determination of nuclear grade and histologic subtype.[1] Yet, the histopathologic analysis of RCC is subject to a high degree of interobserver variability. Multi-institutional studies demonstrate poor to fair agreement between pathologists when determining nuclear grade ($\kappa = 0.22$).[2] When distinguishing histologic subtypes of RCC, agreement is higher between pathologists ($\kappa = 0.43$), but still subject to some interobserver variability.[3] These issues are accentuated during renal mass biopsy (RMB), where undersampling of tumor can limit the ability to grade or subtype RCC in up to 32% of cases.[4]

One potential approach to address the limitations associated with RCC pathologic interpretation is to augment existing workflows with deep learning. Deep learning is a field of artificial intelligence that uses computer algorithms to automatically detect complex patterns in data. Convolutional neural networks (CNNs) are a powerful deep learning tool that has had much success with medical image analysis tasks, such as evaluation of prostate magnetic resonance imagings.[5] In some applications, such as in chest x-ray interpretation or dermatologic lesion evaluation, these CNN-based algorithms can outperform practicing physicians.[6,7] Once adequately trained on a sufficient number of exemplar images, CNNs are able to accomplish image recognition tasks much faster than humans. By training CNNs to recognize and characterize RCC, these algorithms could be potentially used to assist in the interpret RCC histopathologic specimens.

In this study, we sought to develop and validate a CNN that could distinguish benign from malignant renal histopathology, as well as characterize tumor grade and histologic subtypes of RCC.

METHODS

Data Acquisition and Image Generation

Data was obtained from The National Institutes of Health Cancer Genome Atlas Data Portal.[8] Composed of over 20

Financial Disclosure: *The authors declare that they have no relevant financial interests.*

From the Department of Urology, University of Michigan, Ann Arbor, MI; the Department of Pathology, University of Michigan, Ann Arbor, MI; the University of Michigan Rogel Cancer Center, Ann Arbor, MI; the Department of Internal Medicine, University of Michigan, Ann Arbor, MI; and the Department of Computer Science and Engineering, University of Michigan, Ann Arbor, MI

Address correspondence to: Michael Fenstermaker, M.D., M.S., Department of Urology, Michigan Medicine, 1500 E Medical Center Drive, SPC 5330, Ann Arbor, MI 48109.

E-mail: mfenster@med.umich.edu

Submitted: January 26, 2020, accepted (with revisions): May 17, 2020

https://doi.org/10.1016/j.urology.2020.05.094
0090-4295

Table 1. Whole-Mount malignant surgical pathologic characteristics included in analysis (N = 42)

	n	%
*Histologic subtype**		
Clear cell	15	35.6
Papillary	15	35.6
Chromophobe	12	28.7
*Fuhrman grade**		
I	1	2.4
II	15	25.6
III	23	54.8
IV	3	7.2
Pathologic stage		
pT1a	12	28.7
pT1b	10	23.8
pT2a	10	23.8
pT2b	3	7.2
pT3a	5	11.9
pT4	2	4.8

* These frequencies do not necessarily reflect analyzed frequencies, as the algorithm purposefully undersampled or oversampled specimens to better reflect expected population at-large frequencies.

participating institutions, this consortium has collected biospecimens from over 11,000 patients undergoing cancer treatment for one of 33 types of cancer. After specimens undergo a quality assurance check at a centralized laboratory, tumor data, including genetic sequencing and histopathology, are uploaded to the publicly-available dataset. Using the legacy data portal, we downloaded high-resolution digital scans of Hematoxylin and Eosin-stained, formalin-fixed, and paraffin-embedded, whole-mount histopathology. Slides were included for analysis in this study if they were identified on the attached pathology report (read by a single pathologist per participating institution) as either normal kidney parenchyma, clear cell RCC, papillary RCC, or chromophobe RCC. In total, samples from 42 patients were included for analysis (Table 1). Other subtypes, such as oncocytomas, are not including in the Cancer Genome Atlas, so samples were limited to only the above histologic subtypes.

We randomly selected areas of tissue 1024 × 1024 pixels in size at 20× magnification, with each patient contributing multiple tissue sections. Each 1024 × 1024 selection was then further divided into equally sized "patches" of tissue for computational analysis, corresponding to approximately 100 um2 samples (Supplemental Figure 1). The patches were then separated according to their pathology report into different cohorts for training the model in one of 3 tasks: distinguishing benign from malignant specimens, identifying the histologic subtype of RCC, or grading RCC specimens. In total, 3000 normal patches and 12,168 RCC patches were created for testing. The algorithm would randomly sample from these patches in a weighted fashion, such that analyzed cohorts were approximately 80% malignant and 20% benign to simulate the approximately 80%/20% malignant/benign split seen among RMBs. Furthermore, among malignant cases, weighted sampling occurred such that 80% of those analyzed were clear cell, 15% papillary, and 5% chromophobe to approximate true population frequencies.[9] To help the model adapt to variations in slide preparation and increase generalizability to novel specimens, our algorithm also performed a series of random transformations to each image, such as rotation, flipping, and skewing of the image.

CNN Architecture

When applied to a new image, a CNN first convolves this image with a learned filter (Fig. 1a).[10] During convolution, the filter systematically scans the image to determine if a specific pattern, such as edges, curves, etc., are present. A feature map, the results of the dot product between the filter and image, is generated for each filter. In our model, filters of size 3 × 3 with 3 channels (to preserve red, green, and blue color channel information) were used with a stride length of 1. We utilized a total of 6 different convolutional layers, 2 layers of 32 filters, 2 of 64 filters, and 2 layers of 128 filters each.

The next stage is the pooling step (Fig. 1b). With pooling, nearby clusters of pixels are reduced to a single pixel representing the maximum value of nearby pixels. Pooling serves 2 purposes. First, it compresses a large image to reduce computational needs. Pooling also allows the model to generalize to novel examples by averaging out minute details. For our study, the maximum pixel value in a 2 × 2 area was pooled with a stride of 2. A single pooling layers was inserted after every 2 consecutive convolutional layers. Finally, after the final pooling step, the image is flattened from a multidimensional image with a length and width to a vector in preparation for neural network processing (Fig. 1b).

After the image is converted to a vector of pixel values, the image is fed through a several fully connected layers before being passed to the output layer (Fig. 1c). The nodes of each layer accept input values from the previous layer and apply an activation function. The model was composed of a total of 5 fully-connected layers. The first 2 fully connected layers utilized 512 neurons, while the final contained 256 neurons. We used rectified linear units (ReLU) as an activation function. For each training cycle, 40% of the nodes were randomly dropped out from the model to prevent overfitting to the training data. This dropout occurred after the second and third fully-connected layers. The final layer of the model varied depending on its predicted output. To distinguish between benign and malignant pathology, a binary output layer with sigmoid activation function was used. To categorize RCC subtype or grade, a categorical output was calculated using a softmax activation function.

The model learned the network parameters over time by minimizing the amount of loss on the training data with each successive training step. Loss is measured in terms of categorical cross-entropy error between the actual pathology and the network's prediction. We minimized the cross entropy using Adam, a variant of stochastic gradient descent.[11] We found no significant difference in model performance using different hyperparameters, and thus used default Adam parameters ($\alpha = 0.0001$, $\beta_1 = 0.9$, $\beta_2 = 0.999$, $\varepsilon = 10^{-8}$). The model made a final prediction for the presence of malignant histology using a probability threshold of >0.5. Model predictions for grade or RCC subtype were determined by the highest probability subtype predicted by the model. The model iterated over the entire training dataset at most 25 times, known as training "epochs" to optimize the model weights. The model did not see any significant decreases in error beyond 25 epochs. After these epochs, the model calculates a set of predictions on a validation set of an additional previously unseen patches of a similar histologic distribution used during training. The performance of the model on this validation set is used to further tune model parameters and improve performance metrics. Training stopped after performance on the validation set did not improve any further, and the best performance model up to that point was stored.

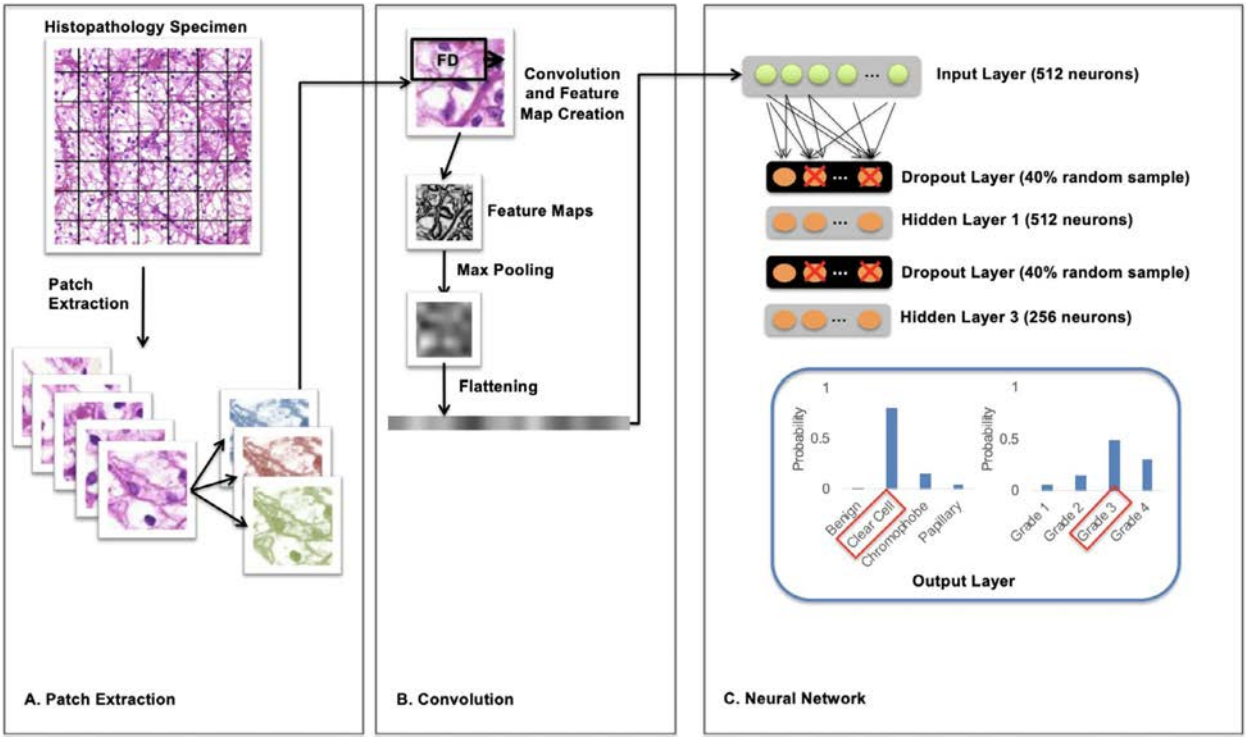

Figure 1. Diagram of convolutional neural network architecture. (A) A digitized H&E slide is split up into 100 um2 patches. Each 2-dimensional patch is represented in memory as a 3-dimensional matrix to preserve red, green, and blue color channel information. (B) A feature detector systematically scans over (convolves) the image searching for a single pattern. Nearby pixels are averaged in the pooling stage to prevent overfitting and reduce the amount of data. The pooled 2D image is flattened to a 1D vector for analysis in a neural network. This flattened image is represented in computer memory as a vector of numbers. (C) The flattened data is fed into a series of fully connected layers of neurons until the final output layer — generating predicted subclass probabilities. The final prediction represents the highest probability prediction. (Color version available online.)

Model Evaluation

We evaluated the model on a final held-out testing dataset of 3486 previously unseen slide patches (3486 RCC, of all grades and histologic subtypes, and 800 normal parenchyma). Again, the model would sample RCC specimens in a weighed fashion to approximate population frequencies of histologic subtypes. Based upon performance of the model with this held out test set, we evaluated final network's accuracy, sensitivity, specificity, positive predictive value, and negative predictive value on this test set. No further model modifications were made after this.

All models were programmed in Python, using the TensorFlow machine learning library.

RESULTS

Characteristics of the renal surgical specimens are tabulated in Table 1. Patient demographic data was mostly redacted in attached The Cancer Genome Atlas pathology reports and was thus omitted. Model learning over time was assessed by diagnostic accuracy on the testing cohort. The model iterated over the initial training and validation sets up to 25 times until we achieved a diagnostic accuracy of 99.98% for distinguishing malignant renal parenchyma from benign, 97.9% categorical accuracy for determining histologic subtype, and 99.6% categorical accuracy for distinguishing tumor grade.

The model was then evaluated in a separate held-out test cohort where the overall diagnostic accuracy for distinguishing benign from malignant parenchyma was 99.4% on a per-100 um2 patch basis at a probability threshold of 0.5. The sensitivity for identifying RCC was 100% and the specificity was 97.1%, with 0 false negative samples and 9 false positive samples. The area under the curve was 0.98.

We then evaluated the model to predict histologic subtype and grade on the held-out out test cohort. Overall categorical accuracy for predicting subtype was 97.5% and grade was 98.4%. Accuracy per subtype is demonstrated in Figure 2.

DISCUSSION

The histopathologic interpretation of RCC is sometimes subject to limitations secondary to inadequate sampling during RMB, as well as variability in interpretation of different histologic subtypes and tumor grades. The utilization of computer-based tools to analyze small foci of disease could help pathologists avoid these limitations. Here, we developed and validated a CNN for rapid and accurate identification for the presence of RCC, the histologic subtype of RCC, and RCC tumor grade. Even when using only single 100 um2 patches of renal tissue, the algorithm was able to correctly identify the associated final whole-mount pathology in greater than 97% of cases. In order for artificial intelligence applications to gain acceptance into clinical practice, they will need to, at the very least, meet or exceed human-level detection of disease.

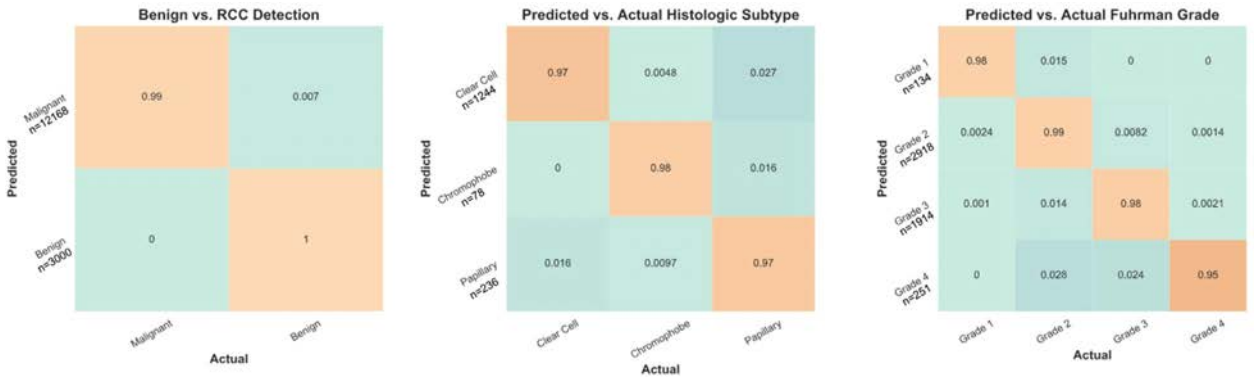

Figure 2. Model predictions on held out test set compared to ground truth. (Color version available online.)

Our results demonstrate that a CNN is capable of identifying the presence of, and characterizing, RCC with a very high degree of accuracy.

We foresee several potential benefits to such an algorithm. Despite the fact that the CNN was trained on whole-mount surgical specimens, the CNN is able to accurately predict final pathology using only small 100 um2 patches of tissue. The low volume of tissue required to make predictions could prove useful in the interpretation of RMB. Detection of small renal masses on cross-sectional imaging is increasingly common, accounting for nearly half of all new RCC diagnoses.[12] Yet, approximately 20% of these small renal lesions are benign, potentially subjecting patients to unnecessary morbidity from surgical excision.[13] RMB is now often employed as a means of distinguishing benign from malignant lesions in order to assist in surgical decision-making. However, there are several limitations associated with RMB. In particular, nondiagnostic biopsies occur in approximately 10%-15% of cases,[14,15] which often results from insufficient tumor sampling.[16] Prior studies have shown that in 12% of RMB specimens, histologic subtyping is not possible, and in 32% of specimens, accurate grading is also not possible.[4] These difficulties often stem from inadequate sampling of tumor, which preclude conclusive or accurate interpretation by pathologists.[16] An algorithm that can predict final pathologic outcomes on the basis of small tissue samples could help in situations where inadequate RMB sampling has occurred.

Beyond this bare-minimum detection capability, we foresee several other potential benefits to computer-assisted diagnosis with machine learning algorithms. Studies suggest that there is fair agreement between pathologists in their interpretation of renal mass biopsies[17]; however, further reductions in interobserver variability could potentially reduce the need for repeat renal biopsies to confirm a diagnosis or reduce the utilization of outside pathologic consultations. This variability is not limited only to RMB; significant variations in interpretation between pathologists are also found in breast biopsies,[18] prostate biopsies,[19] and melanoma specimens.[20] Several studies have demonstrated that computer-aided analysis of histopathologic specimens helps to reduce inter- and intraobserver variability in histopathologic analysis.[21] Computer-assisted diagnosis evaluates an entire specimen in an unbiased manner and can help avoid the phenomenon of "satisfaction of search," whereby identification of an initial positive finding reduces the awareness of a physician toward a second finding.[22] Computers are also particularly suited toward repetitive tasks, which are often important for accurate specimen grading, such a nuclei segmentation[23] and mitotic figure detection.[24] Relegating these tasks to algorithms could mitigate pathologist fatigue and assist with making staging and grading decisions.

Despite the promise of these early results, there are a number of important limitations. Machine learning algorithms require large amounts of data to train accurate models. In order to integrate machine learning algorithms into pathological interpretation, health care systems would need to widely adopt digital pathology systems. Although more widespread implementation in large health care systems is being realized,[25] digitization of pathology slides is still largely restricted to remote consultations, education, and research.[26] In addition, these algorithms were trained on small patch samples extracted from whole-mount renal specimens, which provide a large canvas of cells, and subsequent imaging data to work with. RMB is often limited by undersampling due to small core specimens; a small patch sample of a whole-mount slide is not necessarily equivalent to a RMB specimen of similar dimension. Although the Cancer Genome Atlas does not contain many core biopsy specimens, ideally this algorithm should be further validated on a sample of digitized RMB specimens.

These limitations notwithstanding, the present data demonstrates the potential for machine learning approaches to impact pathologic interpretation of renal histology, accurately characterizing RCC specimens from just 100 um2 samples. Until recently, large data volume requirements and computational resource requirements made machine learning algorithms impractical to train. The increased availability of digitized clinical data and cheap, powerful computing has led to progressively sophisticated applications of machine learning. These early results demonstrate that machine learning applications are able to accurately and rapidly interpret complex

medical imaging data such as renal histopathology. With further advancements and prospective validation, it is foreseeable that such algorithms will assist pathologists in reaching more consistently actionable diagnoses in the future.

References

1. Delahunt B, Cheville JC, Martignoni G, et al. The International Society of Urological Pathology (ISUP) grading system for renal cell carcinoma and other prognostic parameters. *Am J Surg Pathol*. 2013;37:1490–1504.
2. Lang H, Lindner V, de Fromont M, et al. Multicenter determination of optimal interobserver agreement using the Fuhrman grading system for renal cell carcinoma. *Cancer*. 2005;103:625–629.
3. Ficarra V, Martignoni G, Galfano A, et al. Prognostic role of the histologic subtypes of renal cell carcinoma after slide revision. *Eur Urol*. 2006;50:786–794.
4. Leveridge MJ, Finelli A, Kachura JR, et al. Outcomes of small renal mass needle core biopsy, nondiagnostic percutaneous biopsy, and the role of repeat biopsy. *Eur Urol*. 2011;60:578–584.
5. YangX, WangZ, LiuC, et al: Joint detection and diagnosis of prostate cancer in multi-parametric MRI based on multimodal convolutional neural networks.; 2017.
6. RajpurkarP, IrvinJ, ZhuK, et al: CheXNet: radiologist-level pneumonia detection on chest X-rays with deep learning. 2017.
7. Esteva A, Kuprel B, Novoa RA, et al. Dermatologist-level classification of skin cancer with deep neural networks. *Nature*. 2017;542:115.
8. Grossman RL, Heath AP, Ferretti V, et al. Toward a shared vision for cancer genomic data. *N Engl J Med*. 2016;375:1109–1112.
9. Prasad SR, Humphrey PA, Catena JR, et al. Common and uncommon histologic subtypes of renal cell carcinoma: imaging spectrum with pathologic correlation. *RadioGraphics*. 2006;26:1795–1806.
10. Lecun Y, Bottou L, Bengio Y, et al. Gradient-based learning applied to document recognition. *Proc IEEE*. 1998;86:2278–2324.
11. Srivastava N, Hinton G, Krizhevsky A, et al. Dropout: asimple way to prevent neural networks from overfitting. *J Mach Learn Res*. 2014;15:1929–1958.
12. Nguyen MM, Gill IS, Ellison LM. The evolving presentation of renal carcinoma in the United States: trends from the Surveillance, Epidemiology, and End Results program. *J Urol*. 2006;176:2397–2400.
13. Frank I, Blute ML, Cheville JC, et al. Solid renal tumors: an analysis of pathological features related to tumor size. *J Urol*. 2003;170:2217–2220.
14. Patel HD, Johnson MH, Pierorazio PM, et al. Diagnostic accuracy and risks of biopsy in the diagnosis of a renal mass suspicious for localized renal cell carcinoma: systematic review of the literature. *J Urol*. 2016;195:1340–1347.
15. Marconi L, Dabestani S, Lam TB, et al. Systematic review and meta-analysis of diagnostic accuracy of percutaneous renal tumour biopsy. *Eur Urol*. 2016;69:660–673.
16. Evans AJ, Delahunt B, Srigley JR. Issues and challenges associated with classifying neoplasms in percutaneous needle biopsies of incidentally found small renal masses. *Semin Diagn Pathol*. 2015;32:184–195.
17. Kümmerlin I, ten Kate F, Smedts F, et al. Core biopsies of renal tumors: astudy on diagnostic accuracy, interobserver, and intraobservervariability. *Eur Urol*. 2008;53:1219–1227.
18. Elmore JG, Longton GM, Carney PA. et al: Diagnostic concordance among pathologists interpreting breast biopsy specimens. *JAMA*. 2015;313:1122–1132.
19. Shah MD, Parwani AV, Zynger DL. Impact of the pathologist on prostate biopsy diagnosis and immunohistochemical stain usage within a single institution. *Am J Clin Pathol*. 2017;148:494–501.
20. Elmore JG, Barnhill RL, Elder DE, et al. Pathologists' diagnosis of invasive melanoma and melanocytic proliferations: observer accuracy and reproducibility study. *BMJ*. 2017;357.
21. Gavrielides MA, Gallas BD, Lenz P, et al. Observer variability in the interpretation of HER2/neuimmunohistochemical expression with unaided and computer-aided digital microscopy. *Arch Pathol Lab Med*. 2011;135:233–242.
22. Fleck MS, Samei E, Mitroff SR. Generalized "Satisfaction of Search": adverse influences on dual-target search accuracy. *J Exp Psychol Appl*. 2010;16.
23. Irshad H, Veillard A, Roux L, et al. Methods for nuclei detection, segmentation, and classification in digital histopathology: a review-current status and future potential. *IEEE Rev Biomed Eng*. 2014;7:97–114.
24. Ciresan DC, Giusti A, Gambardella LM, et al. Mitosis detection in breast cancer histology images using deep neural networks. *Proc Med Image Comput Assist Interv*. 2013:411–418.
25. Hartman D, Pantanowitz L, McHugh J, et al. Enterprise implementation of digital pathology: feasibility, challenges, and opportunities. *J. Digit. Imaging*. 2017;30:555–560.
26. Pantanowitz L, Evans A, Pfeifer J, et al. Review of the current state of whole slide imaging in pathology. *J. Pathol. Inform*. 2011;2:36.

EDITORIAL COMMENT

The use of artificial intelligence (AI) in medicine is steadily increasing. In the field of urology, AI projects are being used to augment screening, diagnosis, and treatment decisions.[1-3]

Renal masses have emerged as a subject area in which deep learning (DL) networks are being leveraged to aid in both pathologic and radiologic interpretation. For example, in imaging, DL has shown utility in identifying malignant and benign lesions on CT and MRI.[4-6]

In pathology, AI may aid in identification and diagnosis, given the high interobserver variability in histopathologic specimens and growing use of renal mass biopsies. Together, DL models have the potential to optimize the process of identification and diagnosis of renal tumors, augmenting the efficiency of radiologists and pathologists. Long-term goals of any DL platform should be to identify patients with "bad actors" — tumors that are at high risk of metastasizing or those associated with known familial renal cancer syndromes of aggressive phenotype (a particular interest for us at the National Cancer Institute). Ultimately, the ability to predict clinical outcomes should be a primary goal of DL cancer research.

In this study, the authors elegantly present a novel, proof-of-concept study on the efficacy of a convolutional neural network (CNN) to aid in the histologic interpretation of renal masses. By analyzing over 15,000 segments of malignant and benign tissue from 42 patients in The Cancer Genome Atlas, the authors created a robust model to differentiate between cancer and benign histology. The model was validated in a group over 3000 segments of malignant and benign tissue, resulting in a model with a sensitivity of 100%, specificity of 97.1%, and area under the curve of 0.98. Further proving its utility, the model predicted histologic subtype and grade with 97.5% and 98.4% accuracy, respectively. Given that size of the segments analyzed were similar to the size of a renal mass biopsy specimen, it is implied that this CNN could aid in biopsy interpretation. It is important to note that while the model was created and trained on thousands of segments, only 42 patients were used in its development, potentially limiting generalizability and reproducibility of this model if applied to a larger, more heterogeneous population of patients. Regardless, the authors note some important benefits in automating the identification of relatively routine cases, including the reduction of interobserver variability and ability to

automate routine tasks needed for specimen grading. Arguably, the largest advantage of any DL model would be the ability to identify the correct histology with only a small segment of tissue, reducing the nondiagnostic rate of renal mass biopsies. Thus, building upon the present study to create a robust model with renal mass biopsies validated with a "ground truth" of surgical specimens would be of great clinical value. We hope future DL projects build upon this foundation to develop and refine CNN models with the ability to predict clinically meaningful endpoints.

Nitin K. Yerram, Mark W. Ball, Urologic Oncology Branch, National Cancer Institute, National Institutes of Health, Bethesda, MD

References

1. Harmon SA, Sanford TH, Brown GT, et al. Multiresolution application of artificial intelligence in digital pathology for prediction of positive lymph nodes from primary tumors in bladder cancer. *JCO Clin Cancer Inform*. 2020;4:367–382.
2. Strom P, Kartasalo K, Olsson H, et al. Artificial intelligence for diagnosis and grading of prostate cancer in biopsies: a population-based, diagnostic study. *Lancet Oncol*. 2020;21:222–232.
3. Yu SH, Kim MS, Chung HS, et al. Early experience with Watson for Oncology: a clinical decision-support system for prostate cancer treatment recommendations. *World J Urol*. 2020.
4. Baghdadi A, Aldhaam NA, Elsayed AS, et al. Automated differentiation of benign renal oncocytoma and chromophobe renal cell carcinoma on computed tomography using deep learning. *BJU Int*. 2020; 125:553–560.
5. Zabihollahy F, Schieda N, Krishna S, et al. Automated classification of solid renal masses on contrast-enhanced computed tomography images using convolutional neural network with decision fusion. *Eur Radiol*. 2020.
6. Zhao Y, Chang M, Wang R, et al. Deep learning based on MRI for differentiation of low- and high-grade in low-stage renal cell carcinoma. *J Magn Reson Imaging*. 2020.

https://doi.org/10.1016/j.urology.2020.05.095
UROLOGY 144: 156–157, 2020.