Development of a Deep Learning Algorithm for Segmentation of Kidney Tumor Imaging

Seyed Raein Hashemi[1,2], Boris Gershman[3*#], Vladimir I. Valtchinov[1,2,4*#]

[*]Denotes equal author contribution.

[#]Denotes corresponding authors.

[1]Center for Evidence Based Imaging, Brigham and Women's Hospital, Boston, MA, USA

[2]Department of Radiology, Brigham and Women's Hospital, Boston, MA, USA

[3]Division of Urologic Surgery, Beth Israel Deaconess Medical Center, Boston, MA

[4]Department of Biomedical Informatics, Harvard Medical School, Boston, MA, USA

Correspondence:

Boris Gershman, MD

Division of Urologic Surgery

Beth Israel Deaconess Medical Center

Boston, MA 02215

Phone (617) 667-3739

e-mail: bgershma@bidmc.harvard.edu

Vladimir Valtchinov, PhD

Department of Radiology

Brigham and Women's Hospital

Boston, MA 02215

Phone (617) 525-9711

e-mail: vvaltchinov@bwh.harvard.edu

Keywords: deep learning; radiology; kidney; kidney tumor; segmentation; densenet; auto context; patch fusion

Manuscript Word Count: 1008

INTRODUCTION

Kidney cancer is the 6th most common cancer in men and 10th in women worldwide, with an estimated 140,000 kidney cancer deaths annually.[1] Small kidney tumors comprise approximately half of all diagnosed renal masses due to increased incidental detection by diagnostic imaging.[2-4] Radiographic imaging not only establishes the diagnosis of a kidney tumor, but imaging features are the primary determinant of clinical management.[5] Indeed, imaging features as reflected in nephrometry scoring systems have been correlated with probability of malignant pathology, tumor grade, and technical complexity of nephron sparing surgery.[6-9]

However, the manual extraction of imaging features is time-consuming and subject to heterogeneity across different nephrometry scoring systems. The introduction of deep learning algorithms for classification of radiographic images represents a potential method to automate the analysis of renal tumor imaging.[10-13] One of the initial tasks in the analysis of cross-sectional kidney imaging is the accurate segmentation of kidney parenchyma from tumor and background tissues. The Kidney Tumor Segmentation (KiTS19) Challenge was organized to provide a forum for the development of such segmentation algorithms.[14] Herein, we developed an algorithm using the KiTS19 Challenge dataset to segment kidney parenchyma, tumor, and background.

METHODS

Dataset and Outcomes

The KiTS19 Challenge comprised a dataset of 300 patients who underwent partial or radical nephrectomy at the University of Minnesota from 2010 to 2018[14], provided under the Creative Commons CC-BY-NC-SA license. The dataset consisted of abdominal computed tomography (CT) scans acquired in the arterial-phase with segmentation masks for 3 classifications: background, kidney parenchyma, and kidney tumor. The dataset was split into a training set of 210 randomly selected cases that were made available on March 15, 2019, and a test set of 90 cases that were made available on July 29, 2019. Images were manually segmented and ground truth pixel-level labels were provided for the training set for background, kidney parenchyma, and kidney tumor. Participants were required to develop an algorithm for pixel-level segmentation of these images and to submit classifications for the test images.

Deep Learning Algorithm

We developed a deep learning algorithm using convolutional neural networks (CNN) as described below to produce pixel-level classifications of background, kidney parenchyma, and kidney tumor. We trained the CNN using the 210 training cases and produced pixel-level classifications for the 90 test cases.

Network Architecture

In standard neural network architectures, both high- and low-level features are retained by using skip-connections between layers resulting in a more precise prediction. Resnets sum the result of the identity function on skip connections,[15] whereas DenseNets

concatenate these outputs while forwarding connections from all preceding layers rather than just a previous layer leading to a significantly improved flow of information.[16]

We developed our network architecture based on a deep 3D FC-DenseNet.[17] This DenseNet style architecture is shown in **Figure 1**.[16] The model is trained on local features in a contracting path concatenated with global features in an expanding path. Therefore, the network can learn high-resolution as well as low-resolution local and global features in 3 dimensions. The depth of the model ensures the use of 5 different 3D image feature resolutions in the final prediction.

An initial 2x2x2 convolutional layer with stride of 2 was used to down-sample and reduce the patch size to 64x64x64 while preserving the larger field of view of 128x128x128 followed by three padded 3x3x3 convolutional layers and five dense blocks with a growth rate of 12. Dimension reduction of 0.5 was applied after transition layers to help reduce the number of input feature maps and improve parameter efficiency. There were skip connections between every layer inside the dense blocks. Batch normalization and ReLU layers were applied after all convolutional layers. Dropout rate of 0.2 was used after 3x3x3 convolutional layers in each dense block. Finally, a 1x1x1 convolution with a sigmoid function was used as the final layer. A focal loss function, which is a generalization of the cross entropy function, was used as the cost function to the network.[18]

Training and Testing Strategies

Two separate models were trained with the dataset from KiTS19 challenge. The first network was trained for the purpose of predicting kidney and tumor masks combined and another network exclusively to predict the tumor. The output of the first model was used as a posterior probability map inspired by an auto-context technique and was applied as an extra channel input to the second model to help with focusing the learning features.[19]

To address the problem of effective receptive field and lower accuracy around patch borders, a patch prediction fusion strategy was used as a post-processing step using a second-order spline weighting function placed at the center of each patch.[17] Furthermore, a soft voting approach was used to fuse the weighted probabilities into a single 3D map. An overlap of 50% was used between each 3D patch, therefore each voxel was predicted based on a vote between 32 predictions with augmentations (180-degree rotations in each dimension).

RESULTS

The 3D FC-DenseNet architecture was trained on 210 training cases as described above. The model was then applied to 90 test cases. Representative segmentation maps for select image slices are illustrated in **Figure 2**.

DISCUSSION

Using the KiTS19 Challenge dataset,[14] we developed a deep learning algorithm for the automated segmentation of cross-sectional kidney imaging. The algorithm demonstrated excellent performance in a held-out set of 90 test cases, similar to that of manual segmentation. The application of automated segmentation for kidney tumor imaging has several potential advantages. First, it allows one to massively scale the process of kidney segmentation in cross-sectional imaging compared to manual segmentation, which is precursor to any analysis that requires segmented tumor volumes.[11, 20-23] Perhaps more importantly, it can serve as an important infrastructural component for methods that aim to predict clinical attributes, such as pathologic features, from cross-sectional imaging.[10] The present results represent foundational work for the more ambitious goal of developing deep learning methods to enhance clinical diagnosis from radiologic images.

CONCLUSIONS

We developed a deep learning algorithm for the automated segmentation of cross-sectional kidney imaging that demonstrated excellent performance and may represent a method to replace manual segmentation of images.

Figure 1: The 3D FC-DenseNet architecture using auto-context kidney feature maps as an extra input channel to predict tumor masks.

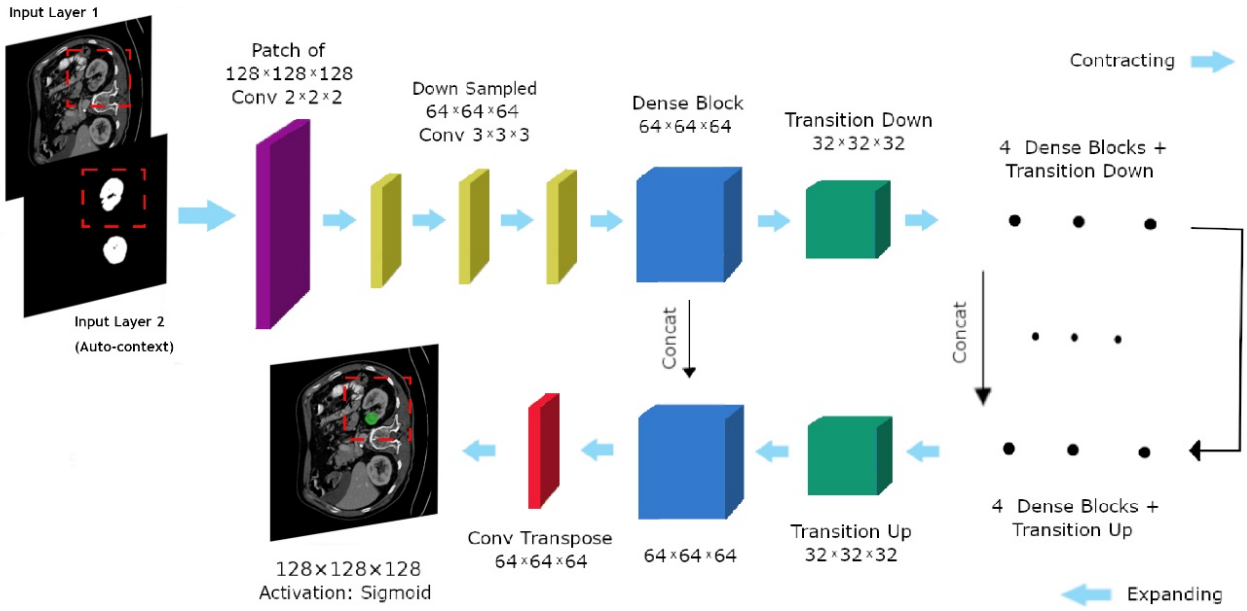
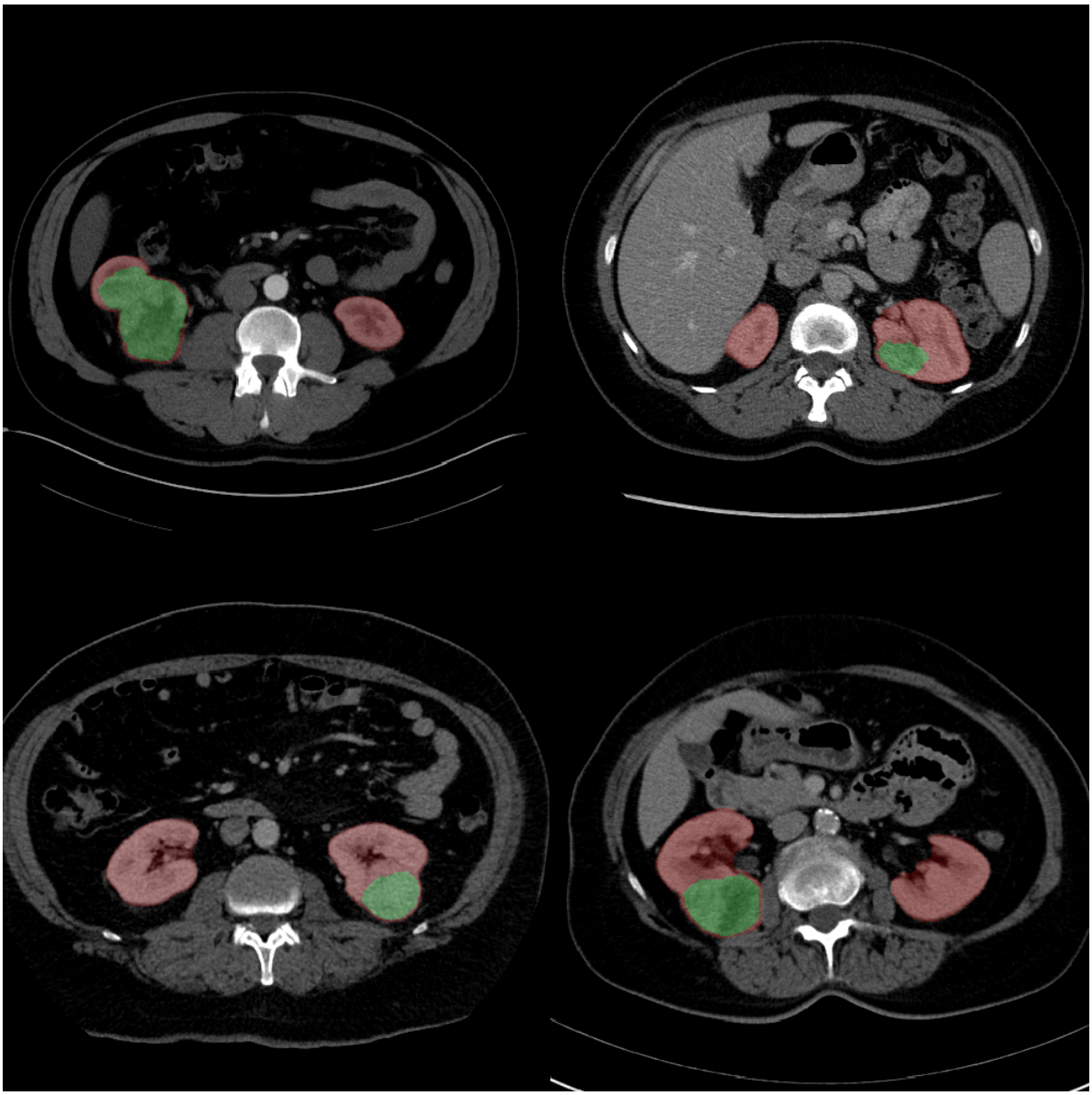


Figure 2: Representative image slices of automated segmentation in select held-out test cases.



ACKNOWLEDGEMENTS

The authors would like to acknowledge the KiTS19 Challenge organizers for availability of the dataset.

CONFLICT OF INTEREST

The authors have no conflicts of interest to disclose.

REFERENCES

1. Capitanio U, Bensalah K, Bex A, Boorjian SA, Bray F, Coleman J, et al. Epidemiology of Renal Cell Carcinoma. Eur Urol. 2019;75(1):74-84.
2. Capitanio U, Montorsi F. Renal cancer. Lancet. 2016;387(10021):894-906.
3. Gill IS, Aron M, Gervais DA, Jewett MA. Clinical practice. Small renal mass. N Engl J Med. 2010;362(7):624-34.
4. Welch HG, Skinner JS, Schroeck FR, Zhou W, Black WC. Regional Variation of Computed Tomographic Imaging in the United States and the Risk of Nephrectomy. JAMA Intern Med. 2018;178(2):221-7.
5. Campbell S, Uzzo RG, Allaf ME, Bass EB, Cadeddu JA, Chang A, et al. Renal Mass and Localized Renal Cancer: AUA Guideline. J Urol. 2017;198(3):520-9.
6. Kutikov A, Uzzo RG. The R.E.N.A.L. nephrometry score: a comprehensive standardized system for quantitating renal tumor size, location and depth. J Urol. 2009;182(3):844-53.
7. Bagrodia A, Harrow B, Liu ZW, Olweny EO, Faddegon S, Yin G, et al. Evaluation of anatomic and morphologic nomogram to predict malignant and high-grade disease in a cohort of patients with small renal masses. Urol Oncol. 2014;32(1):37.e17-23.
8. Osawa T, Hafez KS, Miller DC, Montgomery JS, Morgan TM, Palapattu GS, et al. Comparison of Percutaneous Renal Mass Biopsy and R.E.N.A.L. Nephrometry Score Nomograms for Determining Benign Vs Malignant Disease and Low-risk Vs High-risk Renal Tumors. Urology. 2016;96:87-92.
9. Schiavina R, Novara G, Borghesi M, Ficarra V, Ahlawat R, Moon DA, et al. PADUA and R.E.N.A.L. nephrometry scores correlate with perioperative outcomes of robot-assisted partial nephrectomy: analysis of the Vattikuti Global Quality Initiative in Robotic Urologic Surgery (GQI-RUS) database. BJU Int. 2017;119(3):456-63.
10. Blake P, Sathianathen N, Heller N, Rosenberg J, Rengel Z, Moore K, et al. Automatic R.E.N.A.L. nephrometry scoring using machine learning. European Urology Supplements. 2019;18(1):e904-e5.
11. Taha A, Lo P, Li J, Zhao T. Kid-Net: Convolution Networks for Kidney Vessels Segmentation from CT-Volumes2018.
12. Yala A, Lehman C, Schuster T, Portnoi T, Barzilay R. A Deep Learning Mammography-based Model for Improved Breast Cancer Risk Prediction. Radiology. 2019:182716.
13. LeCun Y, Bengio Y, Hinton G. Deep learning. Nature. 2015;521(7553):436-44.
14. Heller N, Sathianathen N, Kalapara A, Walczak E, Moore K, Kaluzniak H, et al. The KiTS19 Challenge Data: 300 Kidney Tumor Cases with Clinical Context, CT Semantic Segmentations, and Surgical Outcomes. eprint arXiv:190400445. 2019:arXiv:1904.00445.
15. He K, Zhang X, Ren S, Sun J. Deep Residual Learning for Image Recognition. eprint arXiv:151203385. 2015:arXiv:1512.03385.
16. Huang G, Liu Z, van der Maaten L, Weinberger KQ. Densely Connected Convolutional Networks. eprint arXiv:160806993. 2016:arXiv:1608.06993.
17. Hashemi SR, Salehi SSM, Erdogmus D, Prabhu SP, Warfield SK, Gholipour A. Asymmetric Loss Functions and Deep Densely-Connected Networks for Highly-Imbalanced Medical Image Segmentation: Application to Multiple Sclerosis Lesion Detection. IEEE Access. 2018;7:1721-35.
18. Lin TY, Goyal P, Girshick R, He K, Dollar P. Focal loss for dense object detection. IEEE Trans Pattern Anal Mach Intell. 2018.

19. Mohseni Salehi SS, Erdogmus D, Gholipour A. Auto-Context Convolutional Neural Network (Auto-Net) for Brain Extraction in Magnetic Resonance Imaging. IEEE Trans Med Imaging. 2017;36(11):2319-30.
20. Xia KJ, Yin HS, Zhang YD. Deep Semantic Segmentation of Kidney and Space-Occupying Lesion Area Based on SCNN and ResNet Models Combined with SIFT-Flow Algorithm. J Med Syst. 2018;43(1):2.
21. Feng Z, Rong P, Cao P, Zhou Q, Zhu W, Yan Z, et al. Machine learning-based quantitative texture analysis of CT images of small renal masses: Differentiation of angiomyolipoma without visible fat from renal cell carcinoma. Eur Radiol. 2018;28(4):1625-33.
22. Cui EM, Lin F, Li Q, Li RG, Chen XM, Liu ZS, et al. Differentiation of renal angiomyolipoma without visible fat from renal cell carcinoma by machine learning based on whole-tumor computed tomography texture features. Acta Radiol. 2019:284185119830282.
23. Lee H, Hong H, Kim J, Jung DC. Deep feature classification of angiomyolipoma without visible fat and renal cell carcinoma in abdominal contrast-enhanced CT images with texture image patches and hand-crafted feature concatenation. Med Phys. 2018;45(4):1550-61.