NESTED 3D NEURAL NETWORKS FOR KIDNEY AND TUMOR SEGMENTATION

A PREPRINT

Olmo Zavala-Romero,
Department of Radiation Oncology
Sylvester Comprehensive Cancer Center, University of Miami
1475 NW 12th Avenue Miami, FL 33136
osz1@miami.edu

Javier Perez Matos
Electrical and Computer Engineering,
FAMU-FSU, Tallahassee, Florida 32310

Ignacio Alvarez Illan,
Signal Theory and Communications Department
Universidad de Granada, 18071 Granada, Spain
illan@ugr.es

Radka Stoyanova
Department of Radiation Oncology
Sylvester Comprehensive Cancer Center, University of Miami
1475 NW 12th Avenue Miami, FL 33136
rstoyanova@med.miami.edu

Adrian L. Breto
Department of Radiation Oncology
Sylvester Comprehensive Cancer Center, University of Miami
1475 NW 12th Avenue Miami, FL 33136
a.breto@umiami.edu

Isaac Xu
Department of Radiation Oncology
Sylvester Comprehensive Cancer Center, University of Miami
1475 NW 12th Avenue Miami, FL 33136
irx1@miami.edu

Jorge Zavala-Hidalgo
Centro de Ciencias de la Atmósfera
Universidad Nacional Autónoma de México, Mexico City, M'exico
jzavala@atmosfera.unam.mx

Rosario Romero-Centeno
Centro de Ciencias de la Atmósfera
Universidad Nacional Autónoma de México, Mexico City, México
rosario@atmosfera.unam.mx

ABSTRACT

***K*eywords** First keyword · Second keyword · More

1 Introduction

Kidney cancer has an high number of diagnosed cases each year, being invasive techniques the most common treatment[5]. Machine learning techniques have been proven to provide valuable tools in medical imaging, increasing the accuracy in diagnosis for several diseases[2]. In recent years, deep learning techniques have increased the performance of classical machine learning approaches in several imaging problems, including medical image analysis[4]. In this work, we propose a deep learning approach for kidney cancer detection and segmentation, as participants of the Kits19 challenge.

The code will be freely available for others to use on their own data at `https://github.com/naclet/Kits19` after the publication of this work.

2 Methodology

2.1 Data

The data used in this study belongs to the Kits19 challenge[3], which consists of 544 patients who underwent radical nephrectomy of partial nephrectomy, reviewed at the University of Minnesota, with a selection of 300 patients that met the inclusion criteria. These patients have annotated CT abdominal images, with two annotations: kidneys and tumors. The data was separated in training and testing data, and the annotations where only available in the 210 patients training set to the challenge participants.

2.2 Preprocessing

The proposed network architecture, described in section 2.3, receives and image of size 168^3. In order to obtain an image with this dimensions for all our training cases, they had to be preprocessed to a region of interest (ROI) with these dimensions. The proprocessing steps include: a resampling of the image to an isotropic pixel size of $2 \times 2 \times 2$ mm and cropping the CT image to a resolution of 168^3 from the center of the image, if the CT image was smaller, the ROI is padded with zeros. Finallly the ROI is normalized so that the intensity values of all the images range from 0 to 1.

2.3 Neural Network Architecture

The proposed deep learning architecture is 3D V-Net [REF] with fourteen convolutional layers. This network is first trained to learn how to segment the kidneys, after that the weights are saved and the same architecture its used to learn how to segment the tumors. The architecture of this network is shown in figure 1. Each convolutional layer has relu as the activation function and filters size $3 \times 3 \times 3$, with the exception of the last convolutional layer which uses a sigmoid activation function. Each convolutional layer in the decoding phase has batch normalization to improve the generalization of the network. Finally, the dimensions of the input and output images is 168^3.

The network uses stochastic gradient descent as the optimization method, the number of steps per epoch is 100 and the number of steps when computing the metrics on the validation set is 20. The trainig is stopped after 1000 iterations or when the loss function on the validation set do not improves after 90 iterations.

The difference between training the network to learn how to segment kidneys vs tumors, lays on the loss function being used. In both cases, the metric that is minimized is the negative Sorensen-Dice coefficient (DSC) 1 [1]. But, in the case of the training for tumors, the DSC only takes into account the pixels that are inside the kidneys. This helps the network to avoid learning only zero-valued weights.

$$loss = -\frac{2\sum_{i=1}^{N} p_i t_i}{\sum_{i=1}^{N} p_i + \sum_{i=1}^{N} t_i + \varepsilon} \tag{1}$$

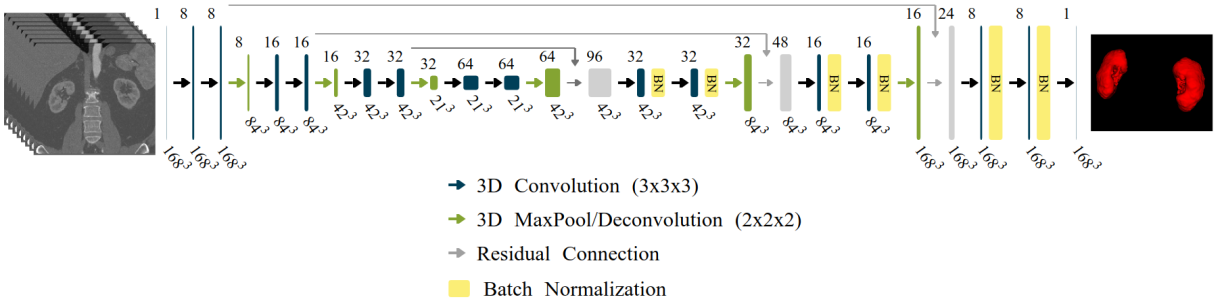

Figure 1: NN architecture

2.4 Posprocessing

For the final classifications, the network is used with the trained weights to predict the locations of the kidneys. During this process the CT image gets preproccessed (resampled, cropped and normalized) and the prediction is made on the obtained ROI. After that, only the top three structures larger than 10 cubic centimeters are kept (if available). Then, the tumors are segmented using the same network but the proper weights for predicting for tumors. The resulting output is masked by the predicted kidneys in the previous step. Finally both segmentaitons are interpolated into its original CT resolution using Nearest Neighbor implementation in Simple ITK.

3 Results

The training dataset was randomly splitted into 90% for training and 10% for validation. We did not have a change to perform 10-fold cross validation before the submission, but will do for the final manuscript. The results obtained by each contour of interest are shown on table 1. The results are separated by the resolution used (ROI and original) and by the dataset used (training, validation, overall).

DataSet	Average DSC	Kidney DSC	Tumor DSC	Kidney ROI DSC	Tumor ROI DSC
Training		0.92 ± 0.07	0.30 ± 0.30	0.93 ± 0.08	0.36 ± 0.35
Validation		0.89 ± 0.06	0.43 ± 0.38	0.89 ± 0.07	0.46 ± 0.40
Overall	0.62 ± 0.17	0.91 ± 0.07	0.32 ± 0.32	0.93 ± 0.08	0.37 ± 0.36

Table 1: Results obtained for the segmentation of kidneys and tumors in the ROI and the orignal resolution.

The results for segmenting the kindeys are much better, with an average DSC of 0.89 ± 0.06 on the validation set, compared with only 0.43 ± 0.38 for the tumors. This shows that finding tumors is a much complicated task that segmenting kindeys.

Figure 2 shows two examples of the obtained kidney and tumor segmentations at the axial and sagittal planes.

4 Conslusions

SUMMARY here

The segmentation of the kidneys can be improved if the network used for training has a larger resolution than 168^3 and the resampled ROI has a smaller pixel size than $2 \times 2 \times 2$ mm. The main reason why this was not tested is the restriction we had in our computer resourcers, for this work four NVIDIA GeForce GTX 1080 were used for the trainnig. These GPUs have 8 GB of memory, and when testing larger models we ran out of memory.

References

[1] Lee R Dice. "Measures of the amount of ecologic association between species". In: *Ecology* 26.3 (1945), pp. 297–302.

[2] Bradley J. Erickson et al. "Machine Learning for Medical Imaging". In: *RadioGraphics* 37.2 (Feb. 2017), pp. 505–515. ISSN: 0271-5333. DOI: 10.1148/rg.2017160130.

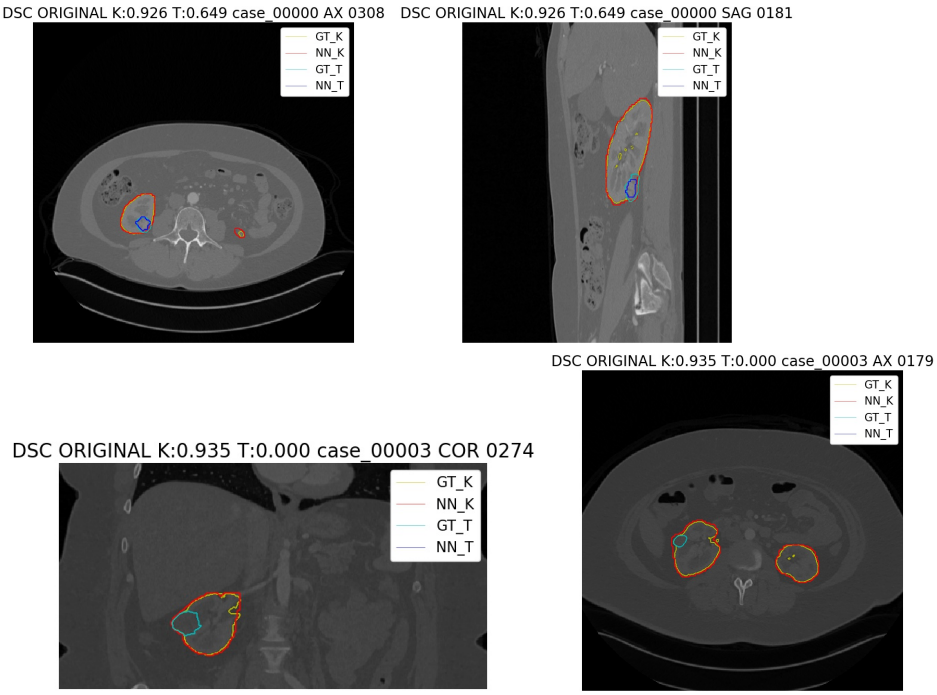

Figure 2: NN architecture

[3] Nicholas Heller et al. "The KiTS19 Challenge Data: 300 Kidney Tumor Cases with Clinical Context, CT Semantic Segmentations, and Surgical Outcomes". en. In: *arXiv:1904.00445 [cs, q-bio, stat]* (Mar. 2019). (Visited on 05/23/2019).

[4] Geert Litjens et al. "A survey on deep learning in medical image analysis". In: *Medical Image Analysis* 42 (Dec. 2017), pp. 60–88.

[5] Maxine Sun et al. "Treatment management of small renal masses in the 21st century: a paradigm shift". eng. In: *Annals of Surgical Oncology* 19.7 (July 2012), pp. 2380–2387. ISSN: 1534-4681. DOI: 10.1245/s10434-012-2247-0.