A Deep Learning Approach for the Segmentation of Kidney, Tumor and Cyst in Computed Tomography Scans

Kartik Narayan Sahoo[1] and Kumaradevan Punithakumar[2(✉)]

[1] Indian Institute of Technology Kharagpur, Kharagpur, India
kartik.sahoo@kgpian.iitkgp.ac.in
[2] Radiology and Diagnostic Imaging, University of Alberta, Edmonton, Canada
punithak@ualberta.ca

Abstract. Kidney cancer, also known as renal cell carcinoma, is a malignant tumor that originates in the kidneys. It is one of the most common types of cancer affecting the urinary system. Kidney tumors can vary in size, location, and aggressiveness, making early detection and accurate diagnosis crucial for effective treatment planning. The proposed method makes use of nnU-Net which is a self-adapting semantic segmentation method, to segment the kidney, tumor and cyst. The proposed neural network model was trained using the datasets provided by the 2023 Kidney and Kidney Tumor Segmentation Challenge hosted by MICCAI 2023 conference. The proposed methodology leveraged the power of deep learning to yield high segmentation accuracy.

Keywords: computed tomography scans · deep convolutional neural network · kidney · tumor · cyst · image segmentation

1 Introduction

Kidney cancer, also known as renal cell carcinoma (RCC), is one of the most prevalent malignancies worldwide, with more than 330,000 new cases being diagnosed annually [6]. The number of cases for kidney tumors have been increasing since the past few decades [2]. It is characterized by the uncontrolled growth of abnormal cells within the kidney. Accurate and precise segmentation of kidney tumors and cysts plays a crucial role in the diagnosis, treatment planning, and monitoring of kidney cancer [7]. In recent years, deep learning techniques, such as the nnU-Net framework [3], have shown remarkable potential in medical image segmentation tasks. nnU-Net is a state-of-the-art deep convolutional neural network architecture that has been successfully applied to various medical imaging tasks. The framework leverages a cascaded U-Net architecture [5], which consists of multiple nested U-Net subnetworks. The network is trained using a combination of dice and cross-entropy loss functions, with extensive data augmentation techniques to enhance robustness and generalization. nnU-Net has demonstrated

remarkable success in various medical imaging applications, including segmentation of organs, tumors, lesions, and abnormalities. Its flexibility, adaptability, and superior performance make it a valuable tool for precise and accurate medical image segmentation tasks. This paper presents an approach for kidney, tumor, and cyst segmentation through deep convolutional neural networks (CNNs), using the nnU-Net architecture. The proposed methodology aims to leverage the power of deep learning to achieve accurate and robust segmentation of the kidney, tumor and cyst structures in medical images, particularly in computed tomography (CT) scans. The proposed neural network utilized for this challenge was trained on a dataset consisting of 489 cases of patients who underwent cryoablation, partial nephrectomy, or radical nephrectomy for suspected renal malignancy. These cases were collected from the years 2010 to 2022 at a M Health Fairview medical center. The CT scan dataset was provided by the 2023 Kidney and Kidney Tumor Segmentation Challenge organizers.

2   Methods

The complete workflow, encompassing both the training and inference stages, is visually illustrated in Fig. 1. Our segmentation approach for kidney, tumor, and cyst regions employed the nnU-Net architecture, without any modifications or adaptations.
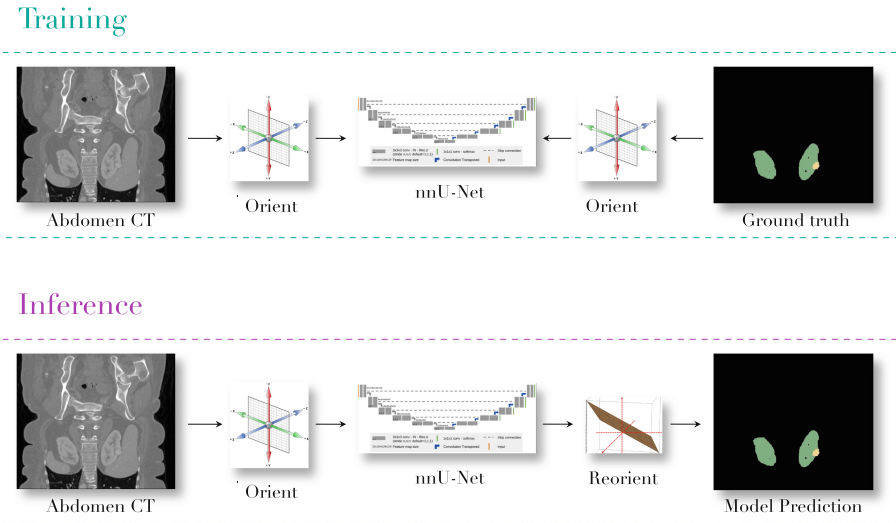

Fig. 1. The proposed neural network based solution to segment the kidney, tumor and cyst from abdomen CT scans.

2.1 Training and Validation Data

Our submission made use of the official KiTS23 training set alone. The dataset is composed of 599 cases with 489 allocated to the training set and 110 in the test set. Only the training set images and ground truths were available, whereas the test set images and ground truths were not revealed to the challenge participants. The challenge training set data (489 cases) was split to 391 training and 98 validation cases for the model. The CT scans are saved as 3D volumes and the dimension range from $(512 \times 512 \times 29)$ to $(512 \times 512 \times 1059)$. The annotated ground truths contain labels comprised of the kidney, tumor and cyst.

2.2 Preprocessing

The dataset's header information containing the position and orientation details of the 3D volume, was removed before the preprocessing step, as we found it to give improved model performance. The preprocessing method involves using the pipeline built with in the nnU-Net architecture. The steps carried out are as follows:

1. **Cropping.** Data undergoes cropping to regions of non-zero values. This cropping process is particularly beneficial as it reduces the size of the data and subsequently minimizes the computational burden.
2. **Resampling.** All data is adjusted to median voxel spacing of the dataset. This ensures uniformity across different scans. Image data is resampled using third-order spline interpolation, which allows for smooth transformations, while the corresponding segmentation masks are resampled using nearest neighbor interpolation to maintain the integrity of the binary segmentation information.
3. **Normalization.** All intensity values within the segmentation masks of the training dataset are collected. The entire dataset is normalized by clipping the intensity values to the 0.5th and 99.5th percentiles of the collected values. This helps to mitigate the impact of outliers. Additionally, a z-score normalization is applied using the mean and standard deviation of all the collected intensity values.
   If the cropping step significantly reduces the average size of patients in the dataset by 1/4 or more in terms of voxels, the normalization is performed only within the mask of nonzero elements and all values outside the mask are set to 0.

2.3 Proposed Method

The model is trained from scratch and evaluated using 5-fold cross validation on the training set. The network uses a combination of dice and cross-entropy loss as the loss function [3].

$$\mathcal{L}_{total} = \mathcal{L}_{dice} + \mathcal{L}_{CE}$$

In our optimization strategy, we employ the Adam optimizer with an initial learning rate of 3×10^{-4} for all experiments. To ensure efficient learning, we monitor the exponential moving average of the training loss. If there is no improvement in this loss for 30 epochs, we adjust the learning rate by reducing it by a factor of 5. If the exponential moving average of the validation loss does not improve by more than 5×10^{-3} within the last 60 epochs and the learning rate drops below 10^{-6}, the training process is stopped.

To prevent overfitting, the nnU-Net performs a variety of data augmentation techniques during training, which includes random rotations, random scaling, random elastic deformations, gamma correction augmentation and mirroring.

To increase the stability of the network, patch sampling is done, where a third of the samples in a batch have atleast one randomly chosen foreground class.

The neural network is trained for 1000 epochs, where an epoch is the iteration over 250 training batches. The training took around 3 days ($\sim$70 h) on the dataset using NVIDIA Tesla A100 (40 GB memory) GPU.

3   Results

The proposed method was quantitatively evaluated over validation CT dataset from over 98 patients. The validation set was derived from the original training set, and the ground truth annotations were available. Evaluation criteria in this research study were based on a method called "Hierarchical Evaluation Classes" (HECs) employed by the organizers. HECs involve combining classes that are subsets of another class to compute metrics for the superset. The HECs used in this study were as follows:

1. **Kidney and Masses**, which included Kidney, Tumor, and Cyst
2. **Kidney Mass**, comprising Tumor and Cyst
3. **Tumor**, focusing solely on Tumor segmentation

Evaluation metrics being used are the Sørensen-Dice and Surface Dice [4]. The class-wise dice scores are shown below:

Table 1 presents the average Sørensen-Dice and Surface Dice values obtained on the validation set of CT scans. The algorithm achieved Sørensen-Dice values of 97.48%, 86.82%, and 84.86% for the kidney and masses, kidney mass, and tumor HECs, respectively. The Surface-Dice values were similar with 96.70%, 77.97% and 73.98% respectively.

Table 2 presents the average Sørensen-Dice and Surface Dice values obtained on the test set of CT scans. The algorithm achieved Sørensen-Dice values of 91.8%, 68.5%, and 60.0% for the kidney and masses, kidney mass, and tumor HECs, respectively. The Surface-Dice values were 84.6%, 53.3% and 45.4% respectively.

The dice and surface-dice score overall were 0.734 and 0.611 respectively.

Table 1. The performance of the proposed algorithm on the validation CT datasets in terms of the Sørensen-Dice and Surface Dice metric. The table reports mean evaluation metrics for each of the HECs on the validation set as defined by the organizers.

Evaluation Metric	Kidney and Masses HEC	Kidney Mass HEC	TumorHEC
Sørensen-Dice (%)	97.485	81.933	78.272
Surface Dice (%)	96.709	77.979	73.989

Table 2. The performance of the proposed algorithm on the test CT datasets in terms of the Sørensen-Dice and Surface Dice metric. The table reports mean evaluation metrics for each of the HECs on the validation set as defined by the organizers.

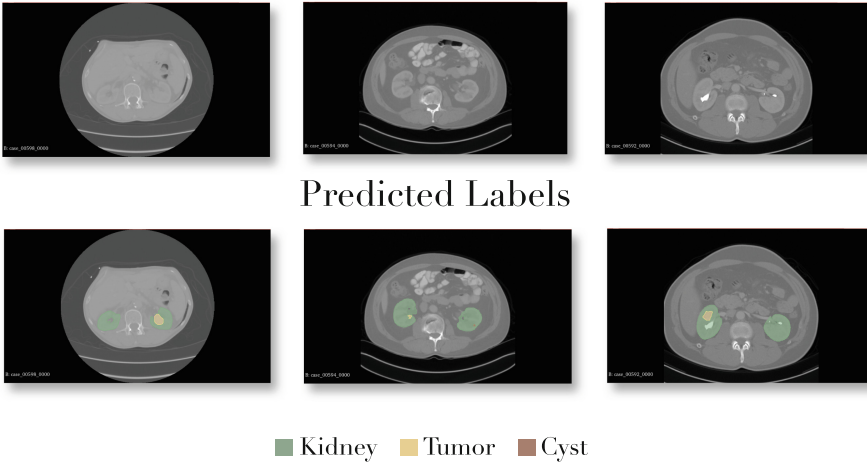| Evaluation Metric	Kidney and Masses HEC	Kidney MassHEC	TumorHEC
Sørensen-Dice (%)	91.8	68.5	60.0
Surface Dice (%)	84.6	53.3	45.4

Fig. 2. Segmentation results on some test set abdomen CT scan images.

4   Conclusion

In this research study, we employed an nnU-Net approach based on deep convolutional neural networks to automatically segment the kidney, tumor, and cyst regions in CT scans. The proposed methodology was evaluated on a validation dataset comprising scans from 98 patients. To assess the performance, we converted the ground truth and predicted images into the three hierarchical evaluation classes (HECs) and employed Deepmind's Surface Distance library for evaluation metrics. The results demonstrated a strong agreement between

the automated predictions and manual delineations, as indicated by Sørensen-Dice coefficient and Surface Dice values. Moving forward, our future work will be directed towards further improving the model's performance, specifically focusing on enhancing the dice score for cyst segmentation. This could be achieved by implementing a nested nnU-Net architecture, utilizing dedicated sub-networks for segmenting each individual component [1].

References

1. Heller, N., et al.: The KiTS21 challenge: automatic segmentation of kidneys, renal tumors, and renal cysts in corticomedullary-phase CT (2023)
2. Hollingsworth, J.M., Miller, D.C., Daignault, S., Hollenbeck, B.K.: Rising incidence of small renal masses: a need to reassess treatment effect. JNCI: J. Natl. Cancer Inst. **98**(18), 1331–1334 (2006). https://doi.org/10.1093/jnci/djj362
3. Isensee, F., et al.: nnU-Net: self-adapting framework for U-Net-based medical image segmentation (2018)
4. Nikolov, S., et al.: Deep learning to achieve clinically applicable segmentation of head and neck anatomy for radiotherapy (2021)
5. Ronneberger, O., Fischer, P., Brox, T.: U-Net: convolutional networks for biomedical image segmentation. In: Navab, N., Hornegger, J., Wells, W., Frangi, A. (eds.) MICCAI 2015. LNCS, vol. 9351, pp. 234–241. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-24574-4_28
6. Scelo, G., Larose, T.L.: Epidemiology and risk factors for kidney cancer. J. Clin. Oncol. **36**(36), 3574–3581 (2018). https://doi.org/10.1200/jco.2018.79.1905
7. Yang, G., et al.: Automatic kidney segmentation in CT images based on multi-atlas image registration. In: 2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, pp. 5538–5541 (2014). https://doi.org/10.1109/EMBC.2014.6944881