Semantic Segmentation of Kidney Tumors Using Variants of U-Net Architecture

https://doi.org/10.3991/ijoe.v18i10.31347

Geethanjali T M(✉), Minavathi, Dinesh M S
P.E.S. College of Engineering, Mandya, Karnataka, India
geethanjalitm@pesce.ac.in

Abstract—Kidney Cancer is one of the most prevalent diseases that is more common in men than women. Detecting kidney tumors at an early stage has been found to increase survival rates of patients. It is therefore important to accurately segment tumors in Computed Tomography(CT) images. To assist radiologist in detection of kidney tumors at early stage, we present a method for segmenting kidney tumors using deep convolutional neural networks. Two U-Net architectures (U-Net, Attention U-Net) variants are used for ensemble training of the models for effective tumor segmentation. Experimental and visual results obtained using the KiTS2019 dataset clearly demonstrate the enhanced Intersection Over Union(IoU) of the ensemble model.

Keywords—Kidney-Tumor segmentation, U-Net architecture, Attention U-Net architecture, ensemble, Computed Tomography, Intersection Over Union

1 Introduction

In recent years, renal cancer is the most commonly reported cancer in adults, ranking ninth in men and fourteenth in women, with obesity, hypertension, smoking and hereditary factors contributing to cancer risk [1,2,3]. Radiologists use CT scans which is commonly used imaging techniques to determine the existence of tumor in kidney organs. However, manually identifying region of interest(ROI) in segmented kidney organs is tedious and time consuming since radiologist must mark ROI in each slice of CT images for each individual patient. The aim of this work is to ensure that radiologists have accurate segmentation approaches to analyse CT images for tumor segmentation to assist them in detecting tumors. Size of kidney and their tumors vary considerably across patients depending on their location and shape. Moreover, tumors may extend beyond the organ and can appear either in a regular shape, distorted or scattered.

There has been considerable interest in applying deep learning architectures to segment kidney tumors in CT images. Few of the published works have used unsupervised methods [4]. Specifically, the aim of the work is to estimate kidney tumor segmentation results by integrating different architectures. Focus is on improving segmentation results of kidney tumors by ensembling different architectures.

The use of deep learning methods for analysing medical images has recently proven to be beneficial. Convolutional neural networks (CNNs) [5, 6, 7, 8, 9, 10, 11] have outlined prevalence over traditional algorithms in an assortment of Computer Vision application. Medical images differ significantly from natural images, despite the availability of deep learning techniques for image segmentation. Pixel-level image marking using Fully Convolutional network (FCN) [12] and U-Net [13] architecture are prominently used to segment medical images semantically.

Prashant Jadiya [14] proposed a three-dimensional U-Net and auto encoder architectures that could segment both kidneys and tumor together. Although the 3D U-Net model was unable to separate tumors alone from kidneys, the auto encoder had better success in separating kidneys from tumors. Utilizing Convolutional neural networks, M.Haghighi et al. [15] have created a strategy of learning kidney division assignments both spatially and time-wise utilizing two, three dimensional inputs at the same time. Pediatric patients with changing degrees of hydronephrosis had their segmentation execution tried on normal kidneys and infected kidneys.The mean dice Coefficient for the normal kidneys was 91.4% while the mean dice coefficient for the diseased kidneys was 83.6%.

A strategy depicted by Luana Batista da Cruz et al. [3] for sectioning CT slices of the kidney employs two CNN and post processing models. The first CNN model AlexNet is used for CT slice categorization, while the second model is utilised to segment kidneys.To identify kidney tumors using fuzzy C-means clustering, Mredhula. L et al. [16] formulated a semi-automated segmentation algorithm that showed promising results. Chenxia Wang et al. [17] created a framework utilizing SC-UNet Associated cascade to segment kidneys and tumors in 3D U-Net design utilizing shape and contextual data. The primary SC-UNet handles rough kidney division, the second one handles smooth kidney and tumor division. The research group Seda Arslan Tuncer et al. [18] has created a modern kidney division strategy that employs the spine as a reference. They compared the connected component labelling (CCL) and the K-means strategies and found that the k-means strategy outflanked the CCL approach.

Many deep learning techniques for segmenting kidneys and tumors have been used in the past. We are currently exploring the possibility of analysing abdominal images using ensemble variants of U-Net models and automatically segmentation of kidney tumors.

2 Dataset with experimental setup

2.1 Dataset

To experiment with the models, the KiTS19 dataset [19] is considered, which includes 210 abdominal CT volumes in the format of Neuroimaging Informatics Technology Initiative (NIFTI). Images and ground truths are given in NIFTI organize with shape consisting of number of slices, height and width. The models are trained with split proportion of 80% for training, 10% for validation and remaining 10% for testing.

2.2 Image preparation and pre processing

The dataset comprises of 210 CT volumes separated into 2D slice images in PNG format, which are then resized to128 X 128 from 512 X 512. Data preparation and pre-processing involves the following steps:

- Step 1: Loading KiTS19 volumes and masks in NIFTI format.
- Step 2: Creating 2D CT image and mask slices from 3D image and mask volumes.
- Step 3: 2D image slices are normalized considering the HU values range of [-1500,2000]
- Step 4: Normalized 2D image slices and mask slices are saved in PNG format.
- Step 5: Ground Truth images are encoded with the values 0,1,2 representing background, tumor and kidney respectively.

2.3 Experimental setup

This Section presents an exploratory setup for segmenting kidney tumors. By taking into consideration the batch size of 8 with a learning rate 1e-4 with respect to the memory restrictions of the graphics card, a unique configuration of U-Net variant designs is utilized to construct the models and are subsequently assessed taking imagenet weights and random weights by initially fine tuning the models for few epochs.

In the experiment, Nvidia Quadro RTX 5000 16GB graphics card is used with Ubuntu 20.04 as the operating system.

3 Proposed approach

Many applications benefit from the use of individual deep learning models, but there is also scope for ensembling [20] several deep learning models to carry out the same task. To segment kidney tumors from Computed Tomography images, the present study ensemble deep learning models based on variants of U-Net architecture using imagenet weights and random weights to create more reliable models with better segmentation results. To obtain a final segmented model, the four models (U-Net, Attention U-Net variant with imagenet and random weights) are stacked together with fine tuning in the intial step of training process. This results in better feature extraction leading to improved performance. Figure 1 is a schematic representation of proposed ensemble architecture.

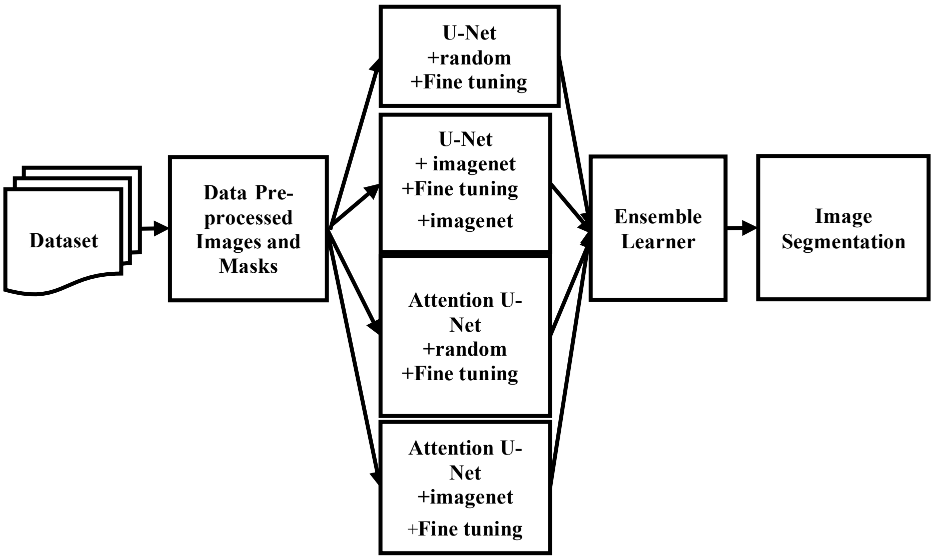


Fig. 1. Proposed model

In Table 1, different segmentation models used for ensembling are presented. All models were written in python using keras and tensorflow.

Table 1. U-Net variants used for experimentation

Segmentation Models
U-Net Variant+ random weights + Model Fine tuning
U-Net Variant+ imagenet weights + Model Fine tuning
Attention U-Net Variant+ random weights + Model Fine tuning
Attention U-Net Variant+imagenet weights + Model Fine tuning

3.1 Segmentation models

This section describes U-Net and Variant U-Net (Attention U-Net) architectures. They allow extracting information about the semantics of an image and describing the specific item within it. The segmentation models are initially fine tuned for a few iterations by training just the decoder without damaging the weights of the encoders. Later all the layers of the segmentation models are released for continuing the training process. Fine tuning is also useful for increasing the performance of segmentation models.

U-Net. Among the convolutional neural networks used in biomedical image segmentation, U-Net is most prominent architecture with few modifications made to accommodate biomedical images. It consists of a path for expanding information, followed by a path for contracting information. Expanding steps consists of upsampling and concatenation followed by convolution operations. The upsampling layer has no weights and doubles the dimension of input. U-Net architecture uses skip connections to simplify the network with few layers. Due to the fact that there are fewer layers to propagate through, the vanishing gradient is reduced, thereby speeding up learning process. As the network learns the feature space, the network restores the skipped layers. Figure 2 represents the U-Net architecture which is fine tuned initially for few iteration used for segmentation of renal tumors in CT images.

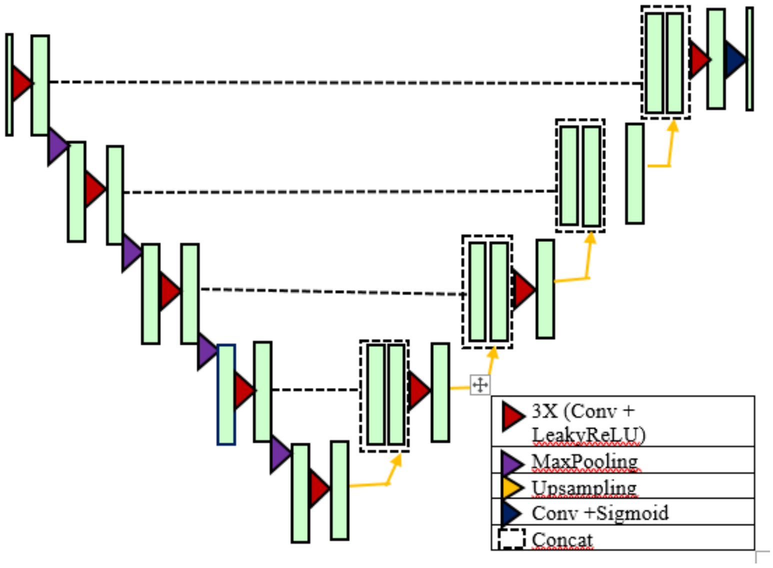


Fig. 2. Proposed U-Net architecture

Attention U-Net. A new U-net architecture, Attention U-Net [21] was designed as an enhancement to the U-Net architecture in 2018.Attention U-Net architecture utilizes only the relevant activations during training, saving computational resources and allowing for better generalization of the model. The architecture employs soft attention at skip connections, which minimizes activations at irrelevant regions while enhancing activations in relevant regions. Skip connections in the architecture stimulate spatial information from both upsampling and down sampling path to retain the good spatial information. Attention gates in turn contribute to the weighting of selected features of interest. Each layer of the proposed architecture is composed of three convolution blocks. During initial training process, the Attention U-Net model is fine tuned for few epochs. Figure 3 potrays the proposed Attention U-Net model to segment kidney tumors.

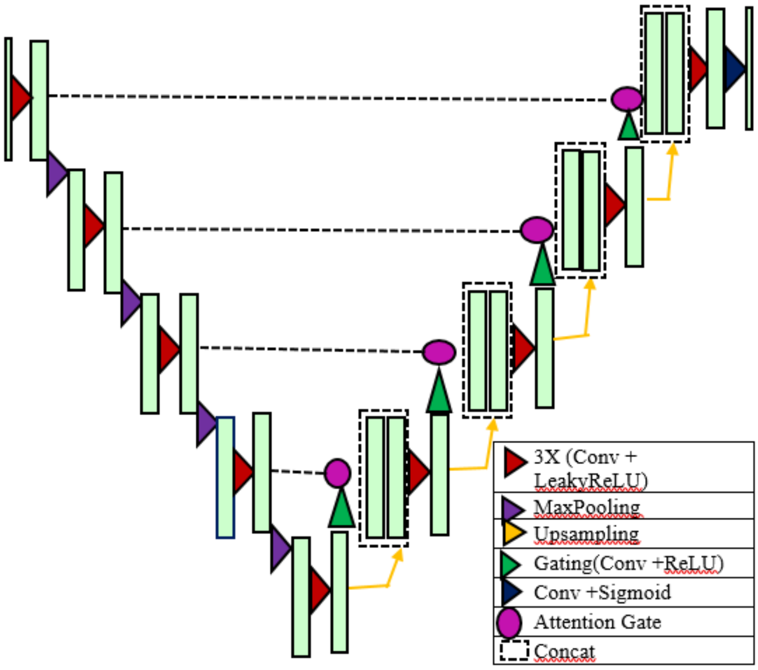


Fig. 3. Proposed attention U-Net model

3.2 Model loss function and evaluation performance metric

For all the models, the loss function used for segmenting multi-class data (background, kidney, tumor) is composed of categorical cross entropy and dice losses. These functions are used to calculate the difference between the labels and predictions. An evaluation performance metric is key to the evaluation of segmentation models. In this work, we apply the Intersection Over Union(IoU) metric. This metric is commonly referred to as the overlapping index, which measures how closely the groundtruth and predicted ouput are overlapping. Collectively the metric in Eq.1 examine the model performance effectively.

$$Intersection\ Over\ Union(IoU) = \frac{True_{positives}}{True_{positives}+False_{positives}+False_{negatives}} \quad (1)$$

3.3 Ensemble model

In order to attain the best segmentation results, the U-Net and Attention U-Net architectures are combined with imagenet and random weights that includes fine tuning approach. Ensembling of the four models is accomplished using weighted averaging, and this in turn results in a better IoU score for kidney tumor.

4 Results and discussion

In this study, as compared to individual models, the ensemble method performs better with hyper parameters being considered the same to train all four models using U-Net and Attention U-Net architectures. IoU scores for kidney tumors are computed taking into account both ground truth and predicted values. Kidney and Tumor mean IoU scores are presented in Table 2 and Table 3. The ensemble model performs well in comparision to the other individual models.

Table 2. Models Kidney mean IoU score

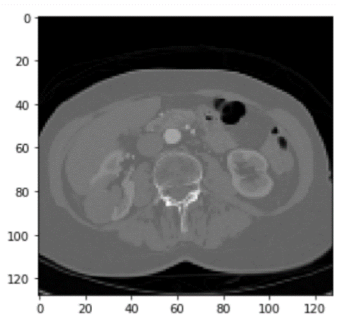
Model	Mean Kidney IoU Score
U-Net (Random Weights)	0.9568
U-Net (ImageNet weights)	0.9549
Attention U-Net (Random weights)	0.9560
Attention U-Net(ImageNet weights)	0.9570
Ensemble Model	**0.9591**

Table 3. .Models Tumor mean IoU score

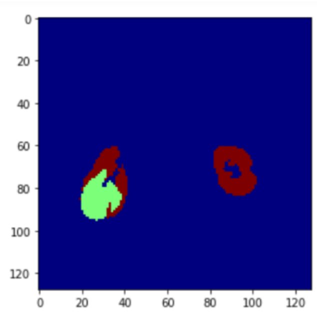
Model	Mean Tumor IoU Score
U-Net (Random Weights)	0.9370
U-Net (ImageNet weights)	0.9344
Attention U-Net (Random weights)	0.9304
Attention U-Net(ImageNet weights)	0.9376
Ensemble Model	**0.9427**

The Figure 4 shows the pixelwise segmentation results obtained from all four models along with ensemble model. The segmentation outputs are visualized almost similar across all the models. However, a closer look reveals ensemble model produces a higher kidney and tumor IoU scores. Under each image, there are two IoU scores pertaining to Kidney and tumor respectively.

Image Slice

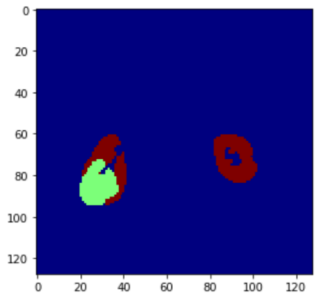


Ground Truth

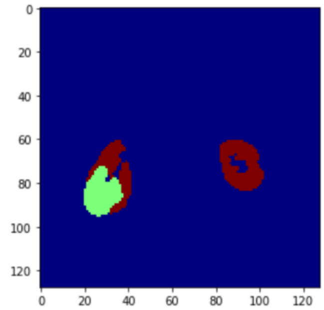


Ensemble Model

U-Net (Random weights)



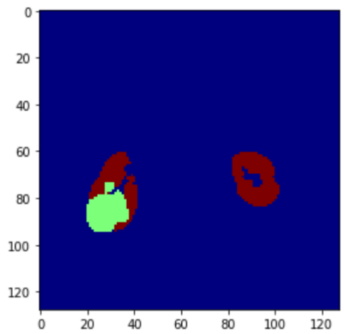
90.84%, 89.72%



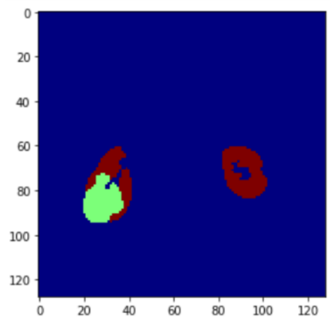
89.37%, 88.05%

U-Net (imagenet weights)

Attention U-Net(Randomweights)

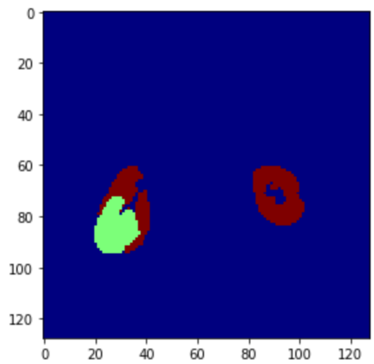


89.33%, 84.53%



90.61%, 88.31%

Attention U-Net (imagenet weights)



90.03%, 87.75%

Fig. 4. Segmentation outcomes of Ensemble model

5 Conclusion and future scope

In order to provide radiologist with a functional method for their decision –making process, this work aims in presenting a method for semantic segmentation of kidney tumors. The current study measures the efficacy of ensemble model for segmenting kidney tumors using KiTS19 dataset of Computed Tomography images. The proposed method utilizes both U-Net and Attention U-Net architectures to segment kidneys and tumors simultaneously. The ensemble model performs better than individual models used in ensembling process in segmenting kidney tumors. Comparing the proposed model to complicated deep learning models with good results in the state of the art, we find it more generalized and simpler. Later on, suitable architectures can be used to segment other organs and to use on different modality images like ultrasound and magnetic Resonance imaging.

6 Acknowledgment

Experimentation of this work was carried out using the equipment funded by the AICTE [Grant No. F.No.9-29/IDC/MODROB/Policy-1/2019-20].

7 References

[1] Cairns P. Renal Cell carcinoma, Canc Biomarkers 2011; 9(1-6):461-73. https://doi.org/10.3233/CBM-2011-0176
[2] S.H. Rossi, T. Klatte, J. Usher-Smith and G.D. Stewart, "Epidemiology and screening for renal cancer," World J. Urol. 36 (9) (2018) 1341–1353, http://dx.doi.org/10. 1007/s00345-018-2286-7J. Clerk Maxwell, A Treatise on Electricity and Magnetism, 3rd ed., vol. 2. Oxford: Clarendon, 1892, pp.68–73. https://doi.org/10.1007/s00345-018-2286-7
[3] Luana Batista da Cru, Jose Denes Lima Araujo, Jonnison Lima Ferreira, Joao Otavio Bandeira Diniz, Aristofanes Correa Silva, Joao Dallyson Sousa de Almei, Anselmo Cardoso de Paiva and Marcelo Gattass, "Kidney segmentation from computed tomography images using deep neural network," Computers in Biology and Medicine, volume 123. https://doi.org/10.1016/j.compbiomed.2020.103906
[4] Fasihi MS, Mikhael WB. Overview of current biomedical image segmentation methods. In: 2016 international conference on computational science and computational intelligence (CSCI). IEEE: 2016. p. 803–8. https://doi.org/10.1109/CSCI.2016.0156
[5] K. He, X. Zhang, S. Ren, J. Sun, Deep residual learning for image recognition, in: Proceedings of the IEEE conference on computer vision and pattern recognition, 2016, pp. 770–778. https://doi.org/10.1109/CVPR.2016.90
[6] G. Huang, Z. Liu, L. Van Der Maaten, K. Q. Weinberger, Densely connected convolutional networks, in: Proceedings of the IEEE conference on computer vision and pattern recognition, 2017, pp. 4700–4708. https://doi.org/10.1109/CVPR.2017.243
[7] A. G. Howard, M. Zhu, B. Chen, D. Kalenichenko, W. Wang, T. Weyand, M. Andreetto, H. Adam, Mobilenets: Efficient convolutional neural networks for mobile vision applications, arXiv preprint arXiv:1704.04861 (2017).

[8] M. Tan, Q. V. Le, Efficientnet: Rethinking model scaling for convolutional neural networks, arXiv preprint arXiv: 1905.11946 (2019).
[9] Jie Hu, Li Shen, Samuel Albanie, Gang Sun, Enhua Wu, Squeeze-and-Excitation Networks, arXiv:1709.01507v4, 16 May 2019.
[10] Christian Szegedy, Sergey Ioffe, Vincent Vanhoucke, Alex Alemi, Inception-v4, Inception-ResNet and the Impact of Residual Connections on Learning, arXiv: 1602.07261v2, 23 August 2016. https://doi.org/10.1609/aaai.v31i1.11231
[11] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna. Rethinking the inception architecture for computer vision. ArXiv preprint arXiv: 1512.00567, 2015. https://doi.org/10.1109/CVPR.2016.308
[12] Jonathan Long, Evan Shelhamer, and Trevor Darrell. Fully convolutional networks for semantic segmentation. In Proceedings of the IEEE conference on computer vision and pattern recognition, pages 3431–3440, 2015. https://doi.org/10.1109/CVPR.2015.7298965
[13] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. Unet: Convolutional networks for biomedical image segmentation. In International Conference on Medical image computing and computer-assisted intervention, pages 234–241. Springer, 2015. https://doi.org/10.1007/978-3-319-24574-4_28
[14] Prashant Jadiya, "Kidney Tumor Segmentation Using Deep Learning," International Journal of Science and Research (IJSR), ISSN: 2319-7064, Volume 8, Issue 12, December 2019, DOI: 10.21275/ART20203792.
[15] M. Haghighi, S.K. Warfield, S. Kurugol, "Automatic renal segmentation in dcemri using convolutional neural networks," 2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018), 2018, pp. 1534–1537. https://doi.org/10.1109/ISBI.2018.8363865
[16] Mredhula. L, M. A. Dorairangaswamy, "DETECTION AND CLASSIFICATION OF TUMORS IN CT IMAGES", Indian Journal of Computer Science and Engineering (IJCSE), ISSN: 0976-5166, Vol. 6 No.2 Apr-May 2015.
[17] Chuanxia Wang, Yuting He, Xiaoming Qi, Ziteng Zhao, Guanyu Yang, Xiaomei Zhu, Shaobo Zhang, Jean-Louis Dillenseger, Jean-Louis Coatrieux, "BiSC-UNet: A fine segmentation framework for kidney and renal tumor," 2019 Kidney Tumor Segmentation Challenge: KiTS19, Miccai 2019, Oct 2019, Shenzen, China, 10.24926/548719.013, hal-02358682. https://doi.org/10.24926/548719.013
[18] Seda Arslan Tuncer, Ahmet Alkan, "Spinal Cord Based Kidney Segmentation Using Connected Component Labeling and K-Means Clustering Algorithm," International Information and Engineering Technology Association(IIETA), Vol. 36, No. 6, December, 2019, pp. 521-527. https://doi.org/10.18280/ts.360607
[19] Nicholas Heller, Niranjan Sathianathen, Arveen Kalapara, EdwardWalczak, Keenan Moore, Heather Kaluzniak, Joel Rosenberg, Paul Blake, Zachary Rengel, Makinna Oestreich , Joshua Dean , Michael Tradewell, Aneri Shah, Resha Tejpaul, Zachary Edgerton, Matthew Peterson, Shaneabbas Raza, Subodh Regmi, Nikolaos Papanikolopoulos and Christopher Weight, "The KiTS19 Challenge Data: 300 Kidney Tumor Cases with Clinical Context, CT Semantic Segmentations, and Surgical Outcomes," ArXiv abs/1904.00445 (2019): n. pag.
[20] Nanditha B R,Geetha Kiran A, Chandrashekar H S, Dinesh M S, Murali S," An Ensemble Deep Neural Network Approach for oral Cancer Screening", International Journal of online and BioMedical Engineering(iJOE), Vol.17, No.02,2021. https://doi.org/10.3991/ijoe.v17i02.19207
[21] Ozan Oktay, Jo Schlemper , Loic Le Folgoc , Matthew Lee , Mattias Heinrich, Kazunari Misawa , Kensaku Mori , Steven McDonagh , Nils Y Hammerla , Bernhard Kainz , Ben Glocker , and Daniel Rueckert, "Attention U-Net: Learning Where to Look for the

Pancreas," 1st Conference on Medical Imaging with Deep Learning (MIDL), Amsterdam, The Netherlands, arXiv:1804.03999v3[cs.CV] , 20 May, 2018.

8 Authors

Geethanjali T M is with P.E.S. College of Engineering, Mandya, Karnataka, India (email: geethanjalitm@pesce.ac.in).

Dr. Minavathi is with P.E.S. College of Engineering, Mandya, Karnataka, India.

Dr. Dinesh M S is with P.E.S. College of Engineering, Mandya, Karnataka, India.

Article submitted 2022-03-31. Resubmitted 2022-05-17. Final acceptance 2022-05-17. Final version published as submitted by the authors.