Original Article

Weight Pruning-UNet: Weight Pruning UNet with Depth-wise Separable Convolutions for Semantic Segmentation of Kidney Tumors

Abstract

Background: Accurate semantic segmentation of kidney tumors in computed tomography (CT) images is difficult because tumors feature varied forms and occasionally, look alike. The KiTs19 challenge sets the groundwork for future advances in kidney tumor segmentation. **Methods:** We present weight pruning (WP)-UNet, a deep network model that is lightweight with a small scale; it involves few parameters with a quick assumption time and a low floating-point computational complexity. **Results:** We trained and evaluated the model with CT images from 210 patients. The findings implied the dominance of our method on the training Dice score (0.98) for the kidney tumor region. The proposed model only uses 1,297,441 parameters and 7.2e floating-point operations, three times lower than those for other network models. **Conclusions:** The results confirm that the proposed architecture is smaller than that of UNet, involves less computational complexity, and yields good accuracy, indicating its potential applicability in kidney tumor imaging.

Keywords: *Depth-wise separable convolution, kidney, kidney tumor segmentation, pruning, weight pruning-UNet*

Submitted: 31-Mar-2021 **Revised:** 10-Jul-2021 **Accepted:** 16-Jul-2021 **Published:** 12-May-2022

Patike Kiran Rao[1], Subarna Chatterjee[2], Sreedhar Sharma[3]

[1]Department of Computer Science and Engineering, MS Ramaiah University of Applied Sciences, [2]Department of Computer Science and Engineering, Faculty of Engineering and Technology, MS Ramaiah University of Applied Sciences, Bengaluru, Karnataka, [3]Department of Nephrology, Kurnool Medical College, Kurnool, Andra Pradesh, India

Address for correspondence:
Mr. Patike Kiran Rao, MS Ramaiah Univeristy of Applied Sciences, Bengaluru, Karnataka, India.
E-mail: kiranraocse@gmail.com

Introduction

The American Cancer Society[1] has reported on the prevalence of kidney cancer in both men and women. Overall, the life-time risk to develop kidney cancer is approximately 1/48 and 1/83 for men and women, respectively. The types of kidney cancer in this study were of an advanced stage. Kidney cancers are generally this advanced stage because the kidneys are situated deep inside the body and are not physically perceived on a physical inspection. Several imaging methods are currently in use to track the growth of kidney tumors. This imaging method has become increasingly popular because it can selectively extract diseased tissues and retain additional stable tissue. This approach was successful in treating small kidney masses. After the precise evaluation of the kidney tumor, details such as the kidney, tumor structure, and others can be collected. In a recent study,[2] it was difficult to derive the essential details from computed tomography (CT) or magnetic resonance imaging scans. Kidney tumors vary in color, form, and scale and have a similar appearance to their parenchyma and other nearby tissues. Given the segmentation of the kidney[3] tumor area, segmenting kidney tumors are extremely difficult.

At present, there is an increased need to deploy deep learning solutions on mobile handheld devices,[4] embedded systems, or machines with minimal resources. An important reason why convolutional neural networks (CNNs) are challenging to train is because they are overparameterized,[5] and they typically require greater computational power and storage space for training and inference. Deep learning researchers have claimed many "pruning" strategies or quantizing learned parameters on broad image datasets.[6-8] Others have concentrated on teaching compact models[9-11] from scratch by factorizing regular convolution layers into depth-wise separable convolution layers for cheaper computations.

Although CNNs have achieved the best results in functional implementations, robustness and accuracy (AC) remain

How to cite this article: Rao PK, Chatterjee S, Sharma S. Weight pruning-UNet: Weight pruning UNet with depth-wise separable convolutions for semantic segmentation of kidney tumors. J Med Signals Sens 2022;12:108-13.

Access this article online
Website: www.jmssjournal.net
DOI: 10.4103/jmss.jmss_108_21
Quick Response Code:

challenging. Ronneberger and Fischer[12] proposed a tool called UNet for automated medical image segmentation to solve these issues. The UNet synthesizes vital information by reducing the cost function in the first half of the network and generates an image in the second half. Inspired by the UNet model, we approached the current challenge of kidney tumor segmentation by proposing a weight pruning (WP)-UNet model. We implemented WP of the UNet with a depth-wise separable convolution architecture, and thus, it refines even tiny regions in the output tumor picture. The system precisely separates the tumor regions of the kidney and offers established quantification and qualitative validity.

Related works

Several computer-aided diagnosis models and artificial neural networks have been developed to classify and segment renal tumors using CT scans. Linguraru *et al*.[13] published a computer-aided method which was used to examine a collection of brain CT scans of 43 patients. In this system, tumors were robustly segmented with approximately 80% overlap. The methodology studied morphological variations between various types of lesions. Lee *et al*.[14] developed a computer program capable of detecting and identifying small renal masses in CT images. Their tests yielded a specific signal-to-noise ratio of 99.63%.

Shah *et al*.[15] presented a segmentation approach using machine learning. Yang *et al*.[16] created a system to automatically segment CT images of the kidney based on multi-atlas registration. First, they recorded a low-resolution image with a series of higher resolution images to create a patient-registered image. Next, the kidney tissues were segmented and aligned to achieve the final segmented production.

Various researchers have also experimented with the segmentation of renal tumors using deep learning. Thong *et al*.[17] used an online patch-wise convolutional kernel to classify the central voxel in two-dimensional (2D) patches. Then, the ConvNet analyzed the CT scan data of each kidney tumor slice.

A Skalski *et al*.[18] demonstrated an efficient hybrid level set approach with elliptical form restrictions for kidney segmentation. The RUSBoost algorithm and decision trees were used to differentiate between kidney and tumor structures, serving as a solution to class imbalance and the need for defining additional voxels. Their model achieved an average precision of 92.1%. Wang *et al*.[19] defined a CNN-based model for kidney segmentation. They proposed a CNN-based segmentation scheme that integrates the bounding box information. They also improved the CNN model by fine-tuning the model for each picture.

Network prototypes

Deep neural networks are superior in their capacity and ability to be generalized. Deep models that learn entirely from data produce excellent results for many tasks when compared with humans. They enhance the plot depth. Researchers have achieved further advances in neural networks. The use of skip links in deep neural networks makes them more trainable to perform tasks such as deep learning. UNet was initially planned to resolve image segmentation, but others such as VGGNet and ResNet were designed for deep classification[13] supervision to further enhance segmentation. Network pruning has been widely studied to compress the CNN models. In early work, network pruning proved to be a valid way to reduce network complexity and overfitting by Hassibi and Stork.[20] Recently, B Hassibi and Stork[20] pruned state-of-the-art CNN models with no AC loss.

Proposed Methods

In this section, we propose the WP-UNet model and describe the modified objective function.

Image preprocessing

All CT images were resized to 256 × 256 pixels in the training set and separated by 255 pixels to normalize the pixel values from 0 to 1.

Dataset

The KiTS19 challenge dataset for kidney tumor disease segmentation was used to assess the performance of WP-UNet. The KiTS dataset[21] consists of 210 high-contrast CT scans collected in the preoperative arterial process. They were chosen from a cohort of subjects who underwent partial or radical nephrectomy[22] for one or more kidney tumors at the University of Minnesota Medical Center and were eligible for inclusion between 2010 and 2018. The volumes included are characterized by different plane resolutions ranging from 0.437 to 1.04 mm, with slice thicknesses ranging from 0.5 mm to 5.0 mm in each case.

The dataset also provides the ground-truth mask of healthy kidney tissue and healthy tumors [Figure 1] for each case. Under the guidance of experienced radiologists, a group of medical students manually generated sample labels with only CT scan image axial projections. A detailed description of the segmentation strategy for the ground truth is described in Heller *et al*.[21] The KiTs challenge dataset is provided with shape (number of slices, height, width) in the standard Neuro Imaging Informatics Technology Initiative format.

Weight pruning-UNet model (proposed architecture)

Figure 2 shows the detailed architecture of the proposed WP-UNet model. The network has the properties of the encoder and decoder structure of the vanilla UNet.[23] As

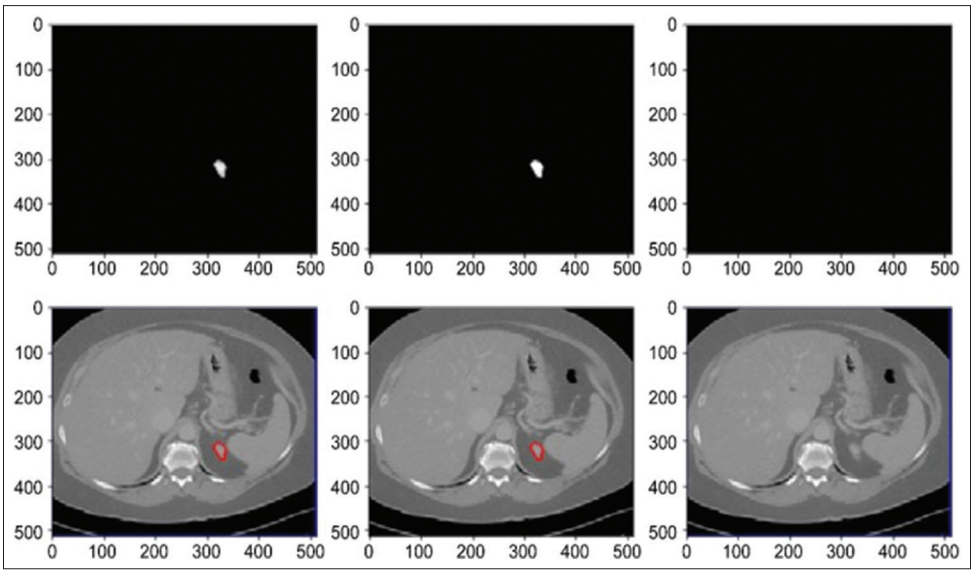

Figure 1: An example of computed tomography scan images from the KiTs19 Challenge dataset.

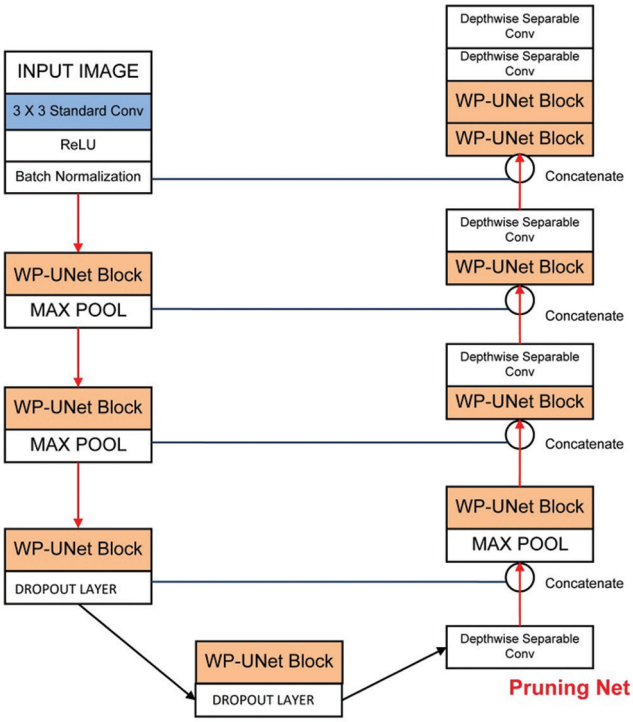

Figure 2: An overview of the detailed architecture of weight pruning-UNet

suggested by Liu *et al*.,[24] first, the input image is passed into the standard convolution layer; subsequently, it is passed to the encoder part of the WP-UNet block. WP-UNet block organized with sequence of layers such as two depth-wise separable convolutional layer, two activation layers, and one batch normalization layer as shown in Figure 3. Here, depth-wise separable convolutional layers are used which is much more commonly used in deep learning (e.g., MobileNet and Xception) for embedded devices.[9]

The proposed model with an input image of size H × W × D, if we do depth-wise separable convolution (stride = 1, padding = 0) with Nc kernels of size e × e × d, where e is even, then the multiplications in transformation for depth-wise separable convolution are (e × e + Nc) × D × (H − e +1) × (W − e + 1) which is less with 2D convolution transformation Nc × e × e × D × (H and e + 1) × (W − e + 1). After training the proposed model, weight-based pruning is applied without compromising the performance of the network. The WP-UNet model uses a weight decay rate of 4e − 5, which has been carefully tuned for the performance on our dataset. In WP-UNet, experiment model includes a dropout layer of rate 0.5 before the up sampling layer. In WP, individual weights in the weight matrix are set to zero. And here to achieve sparsity of S%, we rank the individual weight in weight matrix W according to their magnitude and then set to zero the smallest S%.

Loss function

In this study, the Adam optimizer[25] is applied, which correctly updates the network weights by iteration in the training data. Adam makes an average in the first and second moments of gradients to adapt the learning rate parameter. Sabarinathan *et al*.[26] proposed that the loss function should be the sums of the categorical cross-entropy dice loss channel one (C0) and dice loss channel two (C1), as defined in Eq. (1).

$$\text{Loss} = L + (C_0) + DiceLos\ (C_1) \quad (1)$$

$$Dice\,Loss = 1 - \frac{2\sum_i YiPi + \in}{\sum_i Yi + \sum_i Pi + \in} \quad (2)$$

$$L = -\sum_{j=0}^{M}\sum_{i=0}^{N} Y_{ij} logP_{ij} \quad (3)$$

where L is the cross-entropy loss. In Eq. (2), yi and pi are the ground truth and predicted segmented images,

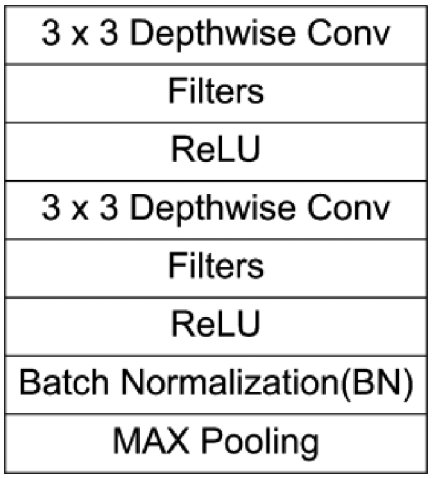

Figure 3: Components of the weight pruning-UNet block

Table 1: Comparison of results between weight pruning-UNet and other models

Model	Training loss	Training accuracy	Mean IOU
UNet	0.5601	97.87	0.435
UNet (depth-wise + BN)	0.4439	93.62	0.362
WP-UNet (network pruning + depth-wise + BN)	0.066	98.43	0.428

BN – Batch normalization; WP – Weight pruning; IOU – Intersection over union

Table 2: Computational comparison between weight pruning-UNet and other models

Model	Parameters	Flops
UNet	5,680,353	62.4e
UNet (depth-wise + BN)	2,601,921	7.8e
WP-UNet (network pruning + depth-wise + BN)	1,297,441	7.2e

BN – Batch normalization; WP – Weight pruning

respectively. Moreover, to ensure the loss function stability, the coefficient ϵ is used.

Performance metrics

The key performance metrics used to measure the WP-UNet performance on the CT scan dataset are explained in this subsection.

Accuracy

AC measures the percentage of correct predictions and is given as,

$$AC = (TP + TN)/(TP + TN + FP + FN) \quad (4)$$

where TP = correctly predicted positive, TN = correctly predicted negative, FP = incorrectly predicted positive, FN = incorrectly predicted negative.

Mean intersection over union

The mean intersection over union (IOU)[20] is a popular evaluation method for semantically segmented images that first determines the IOU for each semantic class and then determines the average over classes. The mean IOU is expressed as follows:

$$\text{Mean IOU} = TP/(TP + FP + FN) \quad (5)$$

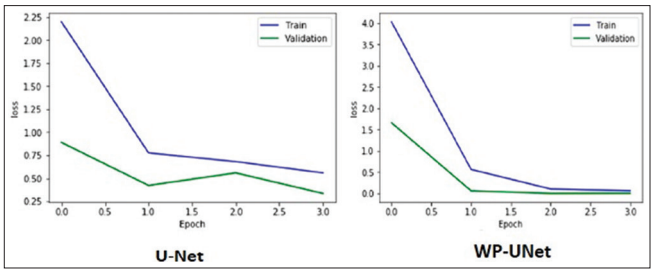

Figure 4: Weight pruning-UNet shows faster converges and better performance during training

Floating-point operations

Floating-point operations (FLOPs) essentially calculate the number of multiplications and additions of floating-point numbers to be performed by the computation device's processor. A neural network in progress requires FLOP calculations to estimate the complexity of the proposed model.

Results

Training

The proposed network was trained with two outputs, namely the kidney and kidney tumor regions. The weight updates were performed using the Adam optimizer with a learning rate of 0.001. The batch size was set to 16, and the total number of epochs was set to a hundred. The training was based on Keras with a Tensorflow backend as a Google Colab deep learning framework enabled with an NVIDIA GPU such as T4 (12 GB memory) with a high memory virtual machine.

The standard dice score is considered an evaluation metric for the performance of the proposed WP-UNet model. We employed 35,865 and 10,158 images as training and validation images, respectively, in our experiments. Table 1 shows the segmentation results of the proposed WP-UNet model for the training and validation images.

From the table, we observe that during training, the proposed method achieves a training AC of 0.98 for the tumor region. Similarly, the computational resource usage of our network is listed in Table 2. Based on the experimental results, we perceive the power of network pruning in the proposed network. Because network pruning is added to the proposed architecture, the total number of flops and parameters is three times smaller than the typical UNet architecture.

As shown in Figure 4, the result of WP-UNet is shown faster convergence with better performance when we

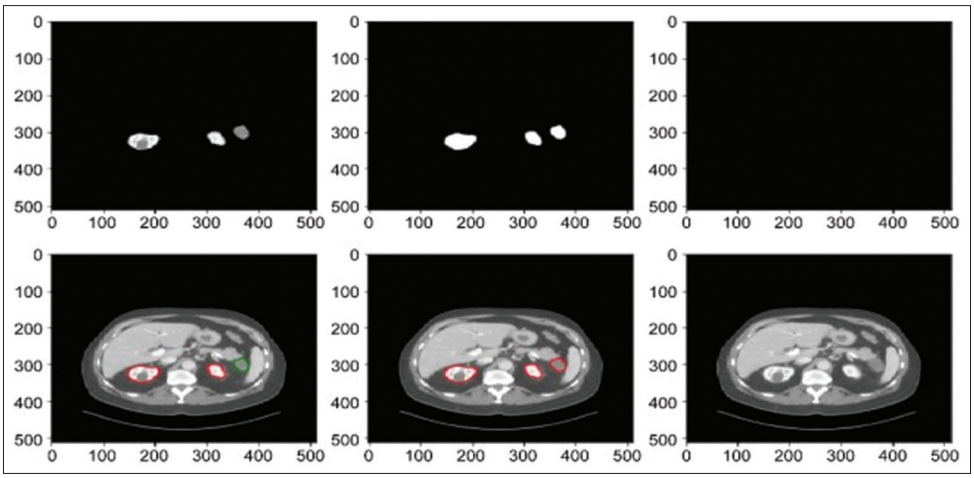

Figure 5: Illustrations of original input computed tomography images and their respective kidney and tumor segmented output images

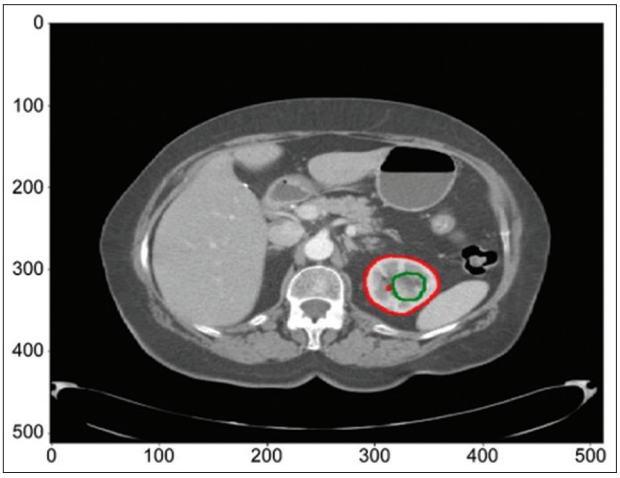

Figure 6: Sample kidney and tumor regions

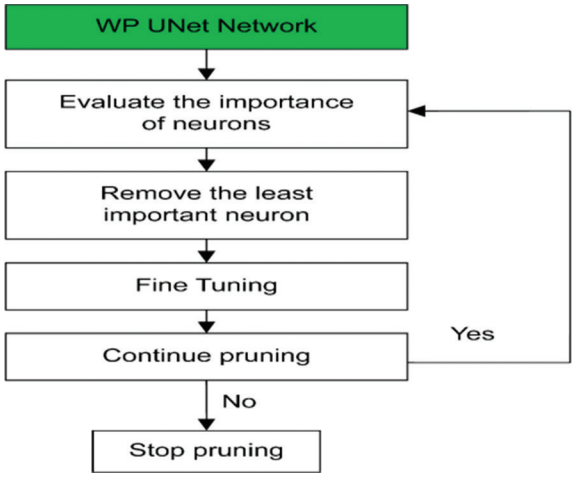

Figure 7: Weight pruning-UNet network pruning

compare with the standard UNet with less number of epochs on large kidney tumor segmentation. In Figure 5, the qualitative effects of the KiTs-19 dataset on the proposed WP-UNet model are shown. We used the provided input images and ground-truth reality images to perform the experiments.

The segmented performance image is depicted in Figure 6. The red colored area is the kidney region in the output picture, and the green-colored part is the kidney tumor. Numerous structures outside the tumor and kidney areas were neglected for simplicity. The final segmented output closely matches the ground truth image from the quantitative results, which demonstrates the usefulness of the proposed WP-UNet.

Conclusions

Medical image segmentation is an important preliminary step in the identification of kidney organ structure and tumor tissues in CT image scans to aid in illness diagnosis, treatment, and general analysis. Early diagnosis is necessary to help in preventing complications that may arise due to late detections. However, with the increasing availability of large biomedical data, the workload on nephrologists, radiologists, and other experts in the field has also increased. To help provide easier, accurate, and timely detections, several deep learning methods have been proposed, most of which have proven to be successful. The UNet architecture is one such model that is widely accepted among researchers for biomedical image segmentation tasks.

In this study, WP-UNet was proposed for the segmentation of kidney tumor data with limited computational resources.

The WP-UNet architecture makes use of depth-wise separable convolutions [Figure 2] and network pruning shown in Figure 7 to reduce the parameters and FLOPs.

Moreover, the WP-UNet deep learning method exhibits a faster inference speed than that of the UNet method. Our findings indicated that the proposed WP-UNet architecture yielded a satisfactory AC. Our system obtained a dice score of 0.9799 and 0.9599 for the preparation and validation sets, respectively. The proposed WP-UNet model achieved

the best segmentation outcomes in terms of the dice score and usage of computational resources. In addition, WP-UNet is shown to have a faster inference speed on test data and is beneficial for situations wherein rapid and accurate segmentation results are required.

Financial support and sponsorship

None.

Conflicts of interest

There are no conflicts of interest.

References

1. American Cancer Society. Available from: https://www.cancer.org/cancer/kidney-cancer/detection-diagnosis-staging/detection.html. [Last accessed on 2019 Dec 17].
2. Hesamian MH, Jia W, He X, Kennedy P. Deep learning techniques for medical image segmentation: Achievements and challenges. J Digit Imaging 2019;32:582-96.
3. Sharma K. Machine Learning Methods for Segmentation in Autosomal Dominant Polycystic Kidney Disease. Munich, Germany: Technische Universität München; 2017.
4. Vaseli H. Designing lightweight deep learning models for echocardiography view classification, 2019, doi: 10.1117/12.2512913.
5. Denil M, Shakibi B, Dinh L, Ranzato M, De Freitas N. Predicting parameters in deep learning. 2013. http://arxiv.org/abs/1306.0543.
6. Lecun Y, Denker JS, Solla SA. Optimal brain damage. Adv Neural Inf Process Syst 1990;2:598-605.
7. Alvarez JM, Salzmann M. Learning the number of neurons in deep networks, Advances in Neural Information Processing Systems, 2016. https://doi.org/10.48550/arXiv.1611.06321.
8. Han S, Mao H, Dally WJ. Deep compression: Compressing deep neural networks with pruning, trained quantization and Huffman coding 2016. http://arxiv.org/abs/1510.00149.
9. Weidendorfer J. Improving landfill monitoring programs Classification of hand movements with the aid of geoelectrical using and geographical information systems Master of Science Thesis: Proc. Asia South Pacific Des. Autom. Conf. ASP-DAC, 2017.
10. Zhang X, Zhou X, Lin M, Sun J. ShuffleNet: An Extremely Efficient Convolutional Neural Network for Mobile Devices. In Proceedings IEEE/CVF Conference on Computer Vision and Pattern Recognition; 2017. p. 6848-56.
11. Qin Z, Zhang Z, Checn X, Wang C, Peng Y. FD-MobileNet: Improved MobileNet with a Fast- Downsampling Strategy. 25th IEEE International Conference on Image Processing (ICIP); 2018. p. 1363-7.
12. Ronneberger O, Fischer P. Thomas Brox: U-Net: Convolutional Networks for Biomedical Image Segmentation. MICCAI; 2015. doi: 10.1007/978-3-319-24574-4_28.
13. Linguraru MG, Wang S, Shah F, Gautam R, Peterson J, Marston Linehan W, *et al*. Automated noninvasive classification of renal cancer on multiphase CT. Med Phys 2011;38:5738-46.
14. Lee HS, Hong H, Kim J. Detection and Segmentation of Small Renal Masses in Contrast-Enhanced CT Images Using Texture and Context Feature Classification. In IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017); 2017. p. 583-6.
15. Shah B, Sawla C, Bhanushali S, Bhogale P. Kidney tumor segmentation and classification on abdominal CT scans. Int J Comput Appl 2017;164:1-5.
16. Yang G, Gu J, Chen Y, Liu W, Tang L, Shu H, *et al*. Automatic Kidney Segmentation in CT Images Based on Multi-atlas Image Registration In 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society; 2017. p. 5538-41.
17. Thong W, Kadoury S, Piche N, Pal CJ. Convolutional networks for kidney segmentation in contrast-enhanced CT scans. Comput Methods Biomech Biomed Engin 2016;6:277-82.
18. Skalski A, Jakubowski J, Drewniak T. Kidney Tumor Segmentation and Detection on Computed Tomography Data. In IEEE International Conference on Imaging Systems and Techniques (IST); 2016. p. 238-42.
19. Wang G, Li W, Zuluaga MA, Pratt R, Patel PA, Aertsen M, *et al*. Interactive medical image segmentation using deep learning with image-specific fine tuning. IEEE Trans Med Imaging 2018;37:1562-73.
20. Hassibi B, Stork D. Second order derivaties for network prunning: Optimal brain surgeon, *Adv. NIPS5*, 1993.
21. Heller N. *et al*. The state of the art in kidney and kidney tumor segmentation in contrast-enhanced CT imaging: Results of the KiTS19 challenge, Med. Image Anal 2021. doi: 10.1016/j.media.2020.101821.
22. 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;182:844-53.
23. Shen W, Wang X, Wang Y, Bai X, Zhang A. Deepcontour: A Deep Convolutional Feature Learned by Positive-Sharing Loss for Contour Detection. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition; 2015. p. 3982-91.
24. Liu R. An intriguing failing of convolutional neural networks and the Coord Conv solution, 2018. https://doi.org/10.48550/arXiv.1904.00445.
25. Kingma DP, Ba J. Adam. A Method of Stochastic Optimization; 2014.
26. Sabarinathan D, Beham MP, Roomi SM. Hyper Vision Net: Kidney Tumor Segmentation Using Coordinate Convolutional Layer and Attention Unit. In National Conference on Computer Vision, Pattern Recognition, Image Processing, and Graphics. Springer; 2019. p. 609-18.