Article

Comparison of Unet3D Models for Kidney Tumor Segmentation

F. Turk[1]*, M. Luy[2] and N. Barisci[3]

[1] Computer Engineering, Graduate School of Natural and Applied Sciences, Kirikkale University, Turkey
[2] Electrical engineering, Institute of Science and Technology, University of Kirikkale, Turkey
[3] Computer Engineering, Institute of Science and Technology, Gazi University, Turkey
* Correspondence: 1788031003@kku.edu.tr; Tel.: +90-505-706-13-73

Abstract: Worldwide, hundreds of thousands of people are diagnosed with kidney cancer and this disease is more common in developed and industrialized countries. Previously, kidney cancer was known as an elderly disease and was seen in people over a certain age; nowadays it is also seen in younger individuals and it is easier to diagnose thanks to new radiological diagnostic methods. A kidney tumor is a type of cancer that is extremely aggressive and needs surgical treatment rapidly. Today, approximately 30% of patients diagnosed with kidney cancer are unfortunately noticed at the stage of metastatic disease (spread to distant organs). The biggest factor that pushes us to this study is that kidney tumors progress unlike other cancer types with little or no symptoms. Therefore, conducting such studies is extremely important for early diagnosis. In this study, we compare the Unet3D models in order to help people who are dealing with difficulties in the diagnosis of kidney cancer. Unet, Unet+ResNet and Unet++ models were compared for image segmentation.

Keywords: Kidney tumor, renal tumor , Unet3D, Unet+ResNet, Unet++ segmentation

1. Introduction

Renal tumors are the third most common urologic cancer and constitute 2-3% of all adult malignancies. Renal Cell Cancer (RCC) is the most common histological subtype and accounts for approximately 85% of renal tumors [1]. Most renal masses are asymptomatic. However, clinical symptoms such as gross hematuria, side pain and palpable mass are occasionally seen. In most cases, it is detected on abdominal imaging.

Nowadays it is easier to catch asymptomatic, small volume localized renal tumors due to the routine use of imaging methods [2]. This led to a significant increase in the incidence of RCC. Localized kidney cancers are those in which the tumor remains confined to the kidney and does not spread to the body. They can be stage 1 (limited to the kidney smaller than 7 cm by size) or stage 2 (greater than 7 cm and limited to the kidney by size).

Renal tumors are divided into benign and malignant tumors. Although benign growth has nothing to do with cancer, it is believed that it has a link to kidney cancer. It's just false information.

The malignant tumor, which we call kidney cancer, is caused by the uncontrolled proliferation of a number of cells in the kidney, resulting in an abnormal mass in the kidney.

The most common causes of kidney cancer are:

- Smoking

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 26 January 2020 doi:10.20944/preprints202001.0314.v1

- Obesity
- Hypertension (high blood pressure patients)

Diagnosis is usually made by whole abdominal ultrasonography (CTU) or abdominal CT (computed tomography) or abdominal MRI performed in patients with other complaints [3]. Some patients are diagnosed before cancer metastasizes to other parts of the body, while in other patients metastatic spread occurs when the cancer is diagnosed for the first time. Kidney cancer can completely be cured if early diagnosed and appropriately treated.

2. Kidney

In a normal person, the kidneys are located on the right and left of the body, towards the back of the rib cage. There is an adrenal gland at the top of each kidney. In a normal person, there are two kidneys, but it may be possible for some people to live with a single kidney. The kidneys are covered with protective adipose tissue called the renal capsule., which is flexible, and if there is a tumor in the kidney, the membrane is stretched.

If the tumor is diagnosed early, it may remain in the capsule and treatment is possible by surgical removal of the kidney [4]. Figure 1. shows the general structure of the kidney [5].

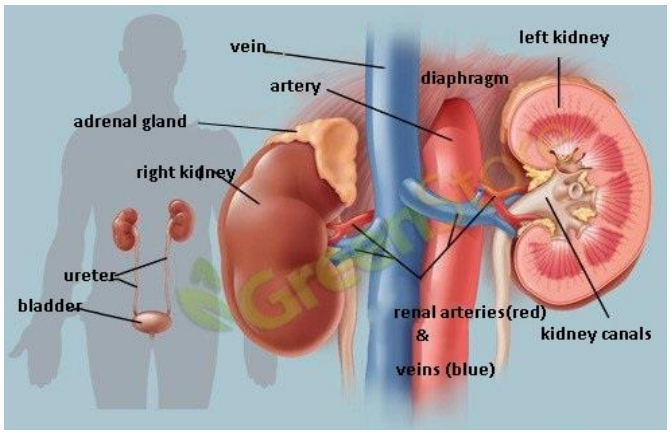

Figure 1. The body Structure of the Kidneys

In recent years, the incidence of kidney cancer in our country and in the world is increasing rapidly. Therefore, kidney cancer surgery operations have also increased. In 2012, 84,400 new cases of kidney cancer were reported in European Union countries. According to the Ministry of Health statistical data for 2013, cases of kidney tumors in Turkey, the number of which was 6534, accounted for 3 percent of all cancer patients [6].

The diagnosis of kidney cancer requires a physical examination of the patient, as well as a careful interpretation of his/her complaints and history The physical examination is important to see whether there is abnormal stiffness or mass in the region matching the kidneys, especially in the abdomen. Blood and urine tests can then be performed. Occasionally, blood in the urine can be caught in patients with kidney tumors. The patient's complaints, history and first evaluation findings of the physician are very important to determine if further research is necessary. In order to understand the condition of the kidney and its surrounding organs, the doctor requests ultrasonography or intravenous urography. Ultrasonography is very useful to determine if there is a tumor in the kidney. Computed tomography (CT) should be performed if a renal tumor is shown or suspected in the images obtained from these tests.

 doi:10.20944/preprints202001.0314.v1

CT shows the condition of the kidney to help understand whether the mass is a cyst or a tumor and whether it has spread or not and whether there is growth in the lymph glands around the kidney. However, an imaging method that distinguishes whether this tumor is benign or malignant (cancer) is not yet available.

The available statistical data suggest that the probability of kidney cancer detected by imaging methods is around 80-85%. Therefore, any mass detected in the kidney should be considered as cancer until further investigation is performed. A needle biopsy may be required to be sure.

To determine the stage of the disease, a lung x-ray and sometimes bone scan can be performed. Figure 2. shows the anatomical view of an early stage renal tumor. Figure 3. shows the anatomical view of a kidney tumor in the advanced stage [7].

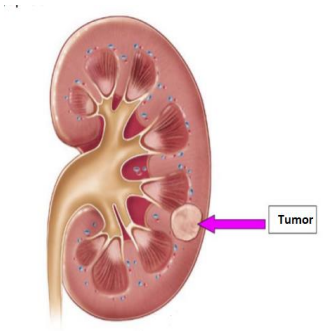

Figure 2. The anatomical view of an early stage renal tumor

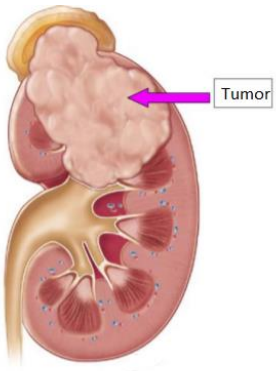

Figure 3. The anatomical view of an advanced renal tumor

3. Materials and Methods

3.1. Dataset

This study employed the Kits19 data set prepared by experienced urologist oncologist writer Christopher Weight and his students specialized in kidney tumors. After the clinical characteristics and imaging data of the patients were collected, manual identification of the renal and tumor margins was performed. The data of a total of 209 patients were drawn by the manual segmentation method. The number of slices showing the kidneys was adjusted to approximately 50 per patient [8]. Figure 4. shows a kidney drawn by manual segmentation in the data set. Data set access is open to the public and can be downloaded via the GitHub page [9].

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 26 January 2020 doi:10.20944/preprints202001.0314.v1

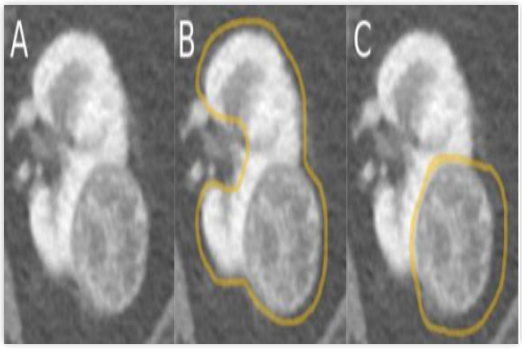

Figure 4. Left: Axial section of the kidney and tumor in the database, Middle: Kidney capsule and tumor, Right: Tumor-only tissue

3.2. Unet Architecture

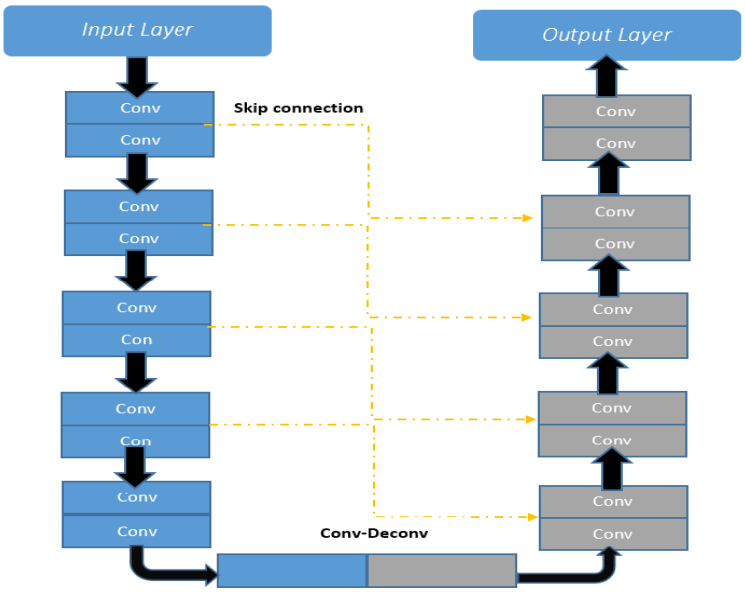

Figure 5. Unet Architecture

Figure 5. shows the classical Unet architecture. Unet was designed by Olaf Ronneberger et al for biomedical image segmentation. The architecture includes two roads. The first path is the contraction path (also called the encoder) used to capture the content in the image. The encoder is a conventional stack of convolution and maximum pooling layers. The second path is the symmetrical expansion path (also called the decoder) used to ensure precise localization using transferred convolutions.

A fully convolutional network (FCN) is one that is trained end-to-end. It contains only convolutional layers and does not contain any dense layers since it can accept images of any size [10, 11]. Our Unet model and layer structures are shown in Table1.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 26 January 2020 doi:10.20944/preprints202001.0314.v1

Table1. Unet model structures

İmage, size	**Operators**
3*3*3*image channel	Input layer
3*3*3*32*32	Conv3d,max pool,
3*3*3*64*64	Conv3d,max pool,
3*3*3*128*128	Conv3d,max pool,
3*3*3*256*256	Conv3d,max pool,
3*3*3*512*512	Conv3d,
3*3*3*256*256	Deconv3d,crop and concat
3*3*3*128*128	Deconv3d,crop and concat
3*3*3*64*64	Deconv3d,crop and concat
3*3*3*32*32	Deconv3d,crop and concat
1*1*1*32	Conv 1*1,sigmoid (output)

3.3. Unet+ResNet Architecture

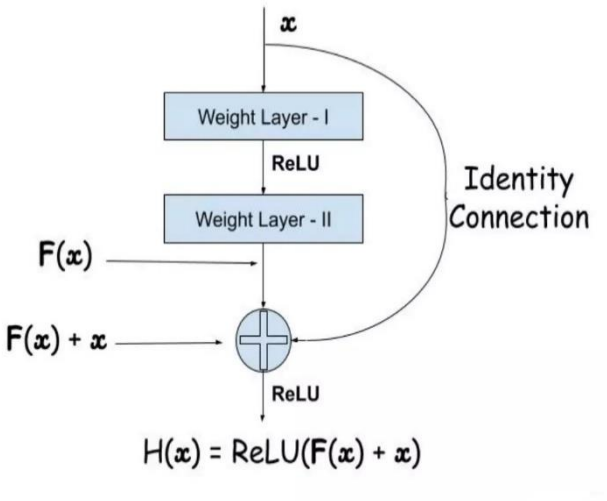

Figure 6. ResNet Architecture

Figure 6. shows the structure of the Resnet architecture. ResNet, unlike its predecessor models, requires adding residual value blocks to the subsequent layers [12]. ResNet distinguishes itself from other classical models with this feature. The residual value changes the system account as follows: [13]. Normally we can map the part from the input to the output with a nonlinear H (x) function. In ResNet architecture, this path is mapped to another nonlinear function defined as F (x): = H (x) - x instead of H (x). In addition, by making a shortcut connection from input to output, the x (input) value is added to the F (x) function arithmetically. Then the function F (x) + x is added together in ReLU. The aim of adding input to the end of layer 2 is to transmit the values from the previous layers to the next layers more strongly. Experiments based on this hypothesis showed that the same function was achieved in both approaches, but the ease of training the network was different [14].

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 26 January 2020 doi:10.20944/preprints202001.0314.v1

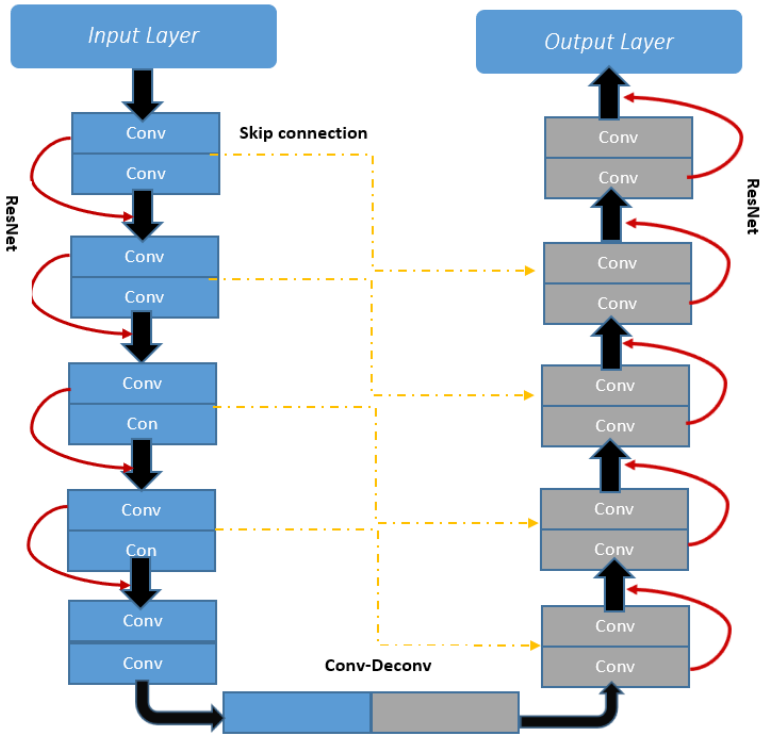

Figure 7. Unet+ResNet Architecture

The Unet+ResNet architecture is shown in figure7. Our Unet+ResNet model and layer structures are shown in Table2. As seen in the model, unlike the Unet model, the layers were transferred to the next layer with the ResNet model.

Table2. Unet+ResNet model structures

İmage, size	**Operators**
3*3*3*image channel	Input layer
3*3*3*32*32	Conv3d,max pool,ResNet
3*3*3*64*64	Conv3d,max pool,ResNet
3*3*3*128*128	Conv3d,max pool,ResNet
3*3*3*256*256	Conv3d,max pool,ResNet
3*3*3*512*512	Conv3d,ResNet
3*3*3*256*256	Deconv3d,crop and concat, ResNet
3*3*3*128*128	Deconv3d,crop and concat, ResNet
3*3*3*64*64	Deconv3d,crop and concat, ResNet
3*3*3*32*32	Deconv3d,crop and concat, ResNet
1*1*1*32	Conv 1*1,sigmoid (output), ResNet

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 26 January 2020 doi:10.20944/preprints202001.0314.v1

3.4. Unet++ Architecture

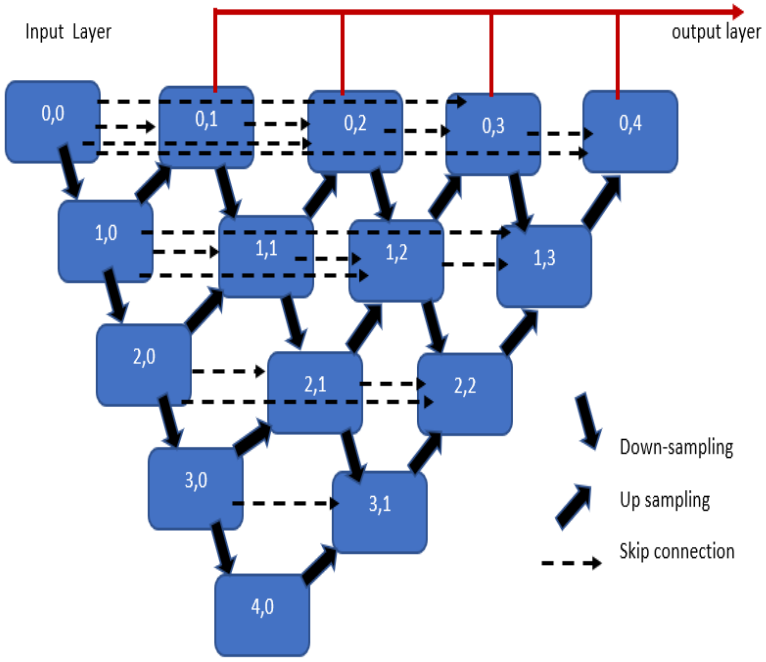

Figure 8. Unet++ Architecture

Unet++ is a new general-purpose image partitioning architecture for more accurate image partitioning. Unet++ consists of an encoder and decoder that are connected through a series of nested dense convolutional blocks. The main idea behind Unet++ is to bridge the semantic gap between the feature maps of the encoder and decoder prior to fusion [15, 16]. Unet++ aims to improve segmentation accuracy by including dense block and convolution layers between the encoder and decoder. Segmentation accuracy is critical for medical images because marginal segmentation errors would lead to unreliable results; thus, will be rejected for clinical settings. Algorithms designed for medical imaging must achieve high performance and accuracy despite having fewer data samples. Acquiring these sample images to train a model can be a resource-consuming process as it requires high quality uncompressed and precisely annotated images vetted by professionals. Table 3 shows the Unet ++ model structure. Unlike other Unet models, it is seen that the convolution layers are made up of successive nested blocks.

Table 3. Unet++ model structures

İmage, size	**Operators**
3*3*3*image channel	Input layer
3*3*3*16*16	Conv3d, downsampling
3*3*3*32*32	Conv3d, downsampling, Nested block
3*3*3*64*64	Conv3d, downsampling, Nested block
3*3*3*128*128	Conv3d, downsampling, Nested block
3*3*3*256*256	Conv3d, downsampling, Nested block
3*3*3*128*128	Deconv3d, crop and concat, Nested block
3*3*3*64*64	Deconv3d, crop and concat, Nested block
3*3*3*32*32	Deconv3d, crop and concat, Nested block
3*3*3*16*16	Deconv3d, crop and concat, Nested block
1*1*1*16	Conv 1*1,sigmoid (output)

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 26 January 2020 doi:10.20944/preprints202001.0314.v1

3.5. Dice Coefficient

In this study, the Dice similarity coefficient (DSC), which is an important criterion for medical segmentation, is taken as reference. The most popular loss function for image segmentation tasks is based primarily on the Dice coefficient, the measure of overlap between the two samples.

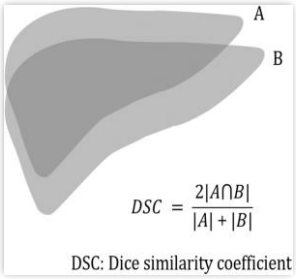

Figure 9. Dice coefficient

Figure 9. shows the dice coefficient and its calculation. This is a measure of the difference between the real face area for segmentation and the area we calculate [17, 18].

4. Results and Discussion

The data of a total of 209 patients from the Kits19 dataset drawn by manual segmentation were made available. We first analyzed the image and calculated slice thickness, window width and position. Later, we created kidney images and masks for these pictures. Finally, we created 3D pictures and masks for the pictures. After converting 3D data to NumPy format, 30 of them were separated for validation, and the remaining datasets were distributed for training. The last 10 datasets were reserved for testing.

Segmentation with Unet3D Models

Unet3D models for renal segmentation was prepared to check the 3D data. The Adam optimizer was used as the optimization algorithm and the ReLU activation function was employed as the activation function. The system was trained for 100000 steps with Nvidia Tesla video card on Tubitak-Truba servers

Kidney Tumor Segmentation with Unet Model

Some of the ground truth images, true masks and predicted masks produced as a result of the validation results are shown in figure 10 for Unet model.

 doi:10.20944/preprints202001.0314.v1

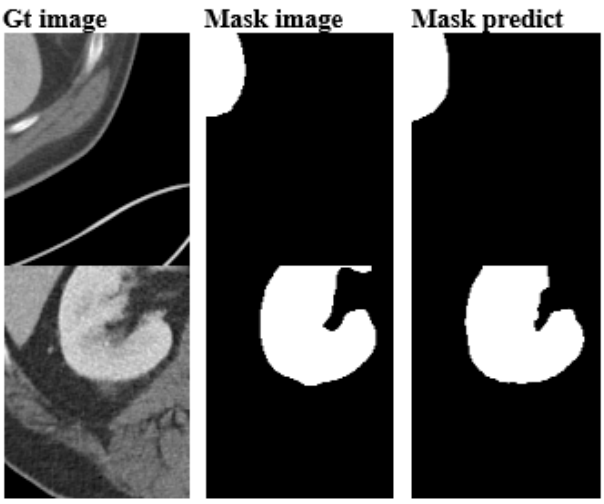

Figure 10. Some cross-sectional images for Unet model

Kidney Tumor Segmentation with Unet + ResNet Model

In contrast to the Unet model, ResNet architecture was carried out with the previous layer in the convolution stage of each layer and the layer was transferred to the next layer in this way.

Some of the ground truth images, true masks and predicted masks produced as a result of the validation results are shown in figure 11 for Unet+ResNet model.

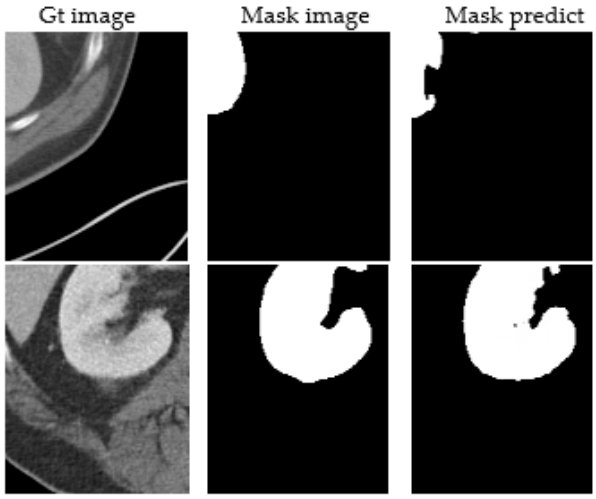

Figure 11. Some cross-sectional images for Unet+Resnet model

Kidney Tumor Segmentation with Unet ++ Model

Unlike the Unet and Unet+Resnet models, each layer was designed to be nested. The output of each layer was computed by means of the sigmoid function. Thus, the calculation of the depth of this model was easier than the other two models.

Some of the ground truth images, true masks and predicted masks produced as a result of the validation results are shown in figure 12 for Unet++ model.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 26 January 2020 doi:10.20944/preprints202001.0314.v1

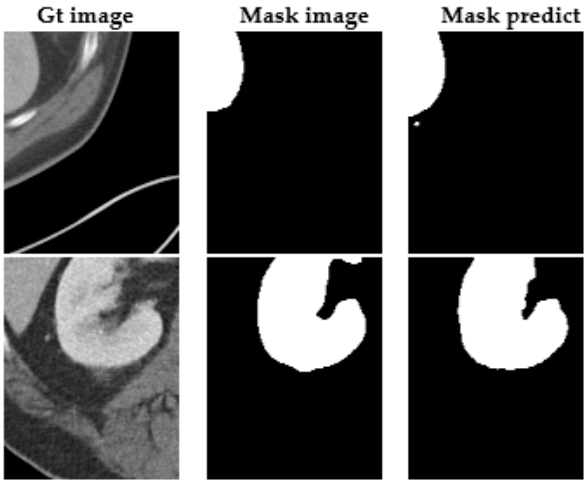

Figure 12. Some cross-sectional images for Unet++ model

Table 4 shows the validation and test results for each of the 3 models. The validation result was 0,807 for the Unet model, 0,871 for the Unet + ReNet model and 0,886 for the Unet ++ model. The test result was 0,931 for the Unet model, 0,946 for the Unet + ResNet model and 0,804 for Unet++ model.

Table 4. Segmentation validation and test results

Model	Validation	Test
Unet	0,807	0,931
Unet+ResNet	0,871	0,946
Unet++	0,886	0,804

Figure 13 shows the test results for all three models. The area selected as the red color appears to be the region in which the Unet ++ model is different in segmentation. Figure 13 shows the ground truth and original masks, as well as the segmentation results for all three models.

 doi:10.20944/preprints202001.0314.v1

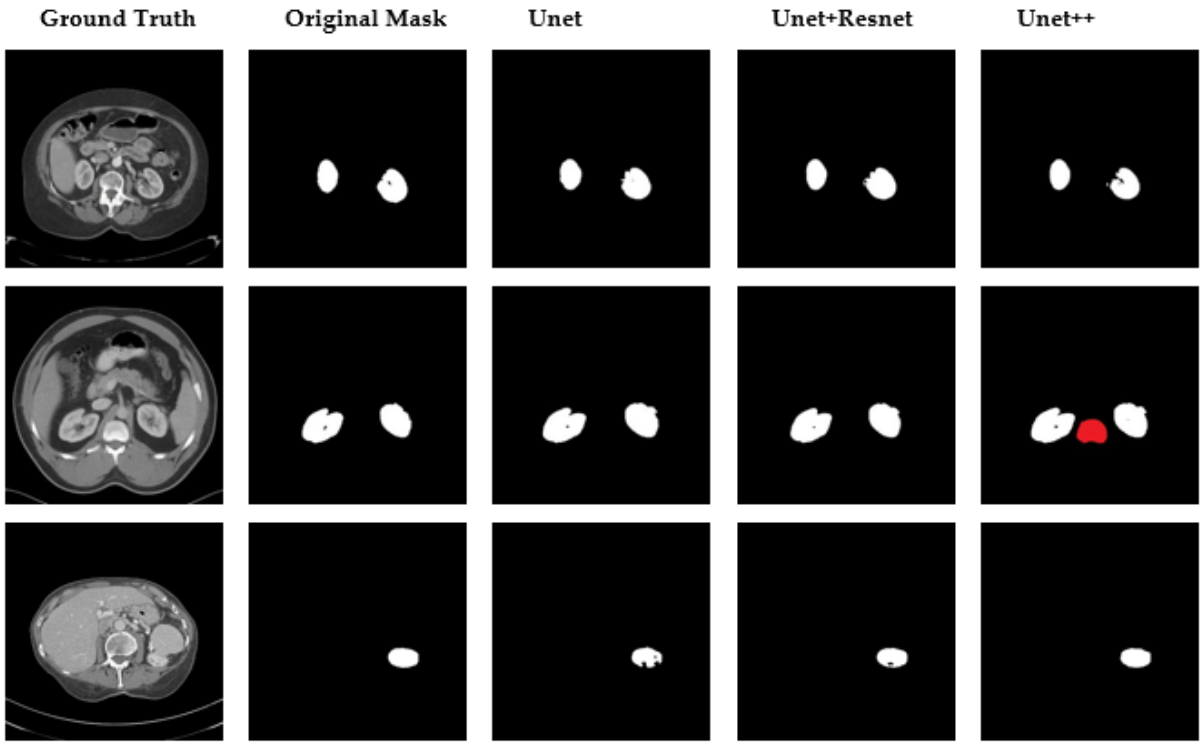

Figure 13. Segmentation results for each model

5 Conclusion

Unet, Unet+ResNet and Unet ++ models were tested under the same conditions using the same parameters. For each designed model, Tubitak-Truba servers of the institution (TUBITAK) were used. The network models were trained for 100000 steps for approximately 8 days at different times.

When the training was completed, the validation DSC with the Unet model was found to be 0,807, 0.871 with the Unet + ResNet model and 0.886 with the Unet ++. The best model weights obtained after 100000 steps were recorded and the actual success of all three models was measured using 10 images that had not initially been trained. Based on the validation results, the Unet ++ model was more successful than the other two models. However, on the test, the Unet + ResNet model produced a significantly higher value than the Unet and Unet ++ models.

Considering all the validation results, the Unet ++ model saves us from the confusion about the number of layers we should not take into account because only the nested layers are taken into account. This adds a new depth to the network model. In addition, this situation has a positive effect on the training results, yielding the highest success levels.

When the test results are taken into consideration, Unet + ResNet model has achieved the highest rate of success. This proves once again how important the ResNet model is, especially in medical image segmentation. Furthermore, the results show that under current conditions for renal tumor segmentation, the Unet-ResNet model may be preferable to the Unet and Unet ++ model. Tumors in

kidneys and other soft tissue organs may, unfortunately, be missed in MR images. Thus, this model might also be more useful for the diagnosis of cancer types like kidney tumors producing no serious symptoms and side effects. By creating a suitable environment, the models can be converted into a hybrid model. Considering the superior aspects of each model, the network's success in validation and test results can be improved. New and fast software for a hybrid model with a high success rate to be used in hospital environments can be created.

Acknowledgment

We would like to thank Tubitak for its services in artificial intelligence.

6 References

1. Parkin, DM; Whelan, SL; Ferlay, J; Raymond, L; Young, J. Cancer incidence in five in continents. v. 7. Lyon: International Agency for Research on Cancer 1997.
2. Chow, WH; Devesa, SS; Warren, JL; et al. Rising incidence of renal cell cancer in the United States. JAMA 1999; 281:1628-31. [CrossRef]
3. Mutlu, D; Volkan, I; Mustafa, ZT. Nephron-Sparıng Surgery in Localızed Kıdney Tumors. Cukurova Journal of Anesthesia and Surgical Sciences, Volume 2, Issue 2. January 2019, 82 – 89.
4. Turk, F; Luy, M; Barisci, N. Machine Learning of Kidney Tumors and Diagnosis and Classification by Deep Learning Methods, International Journal of Engineering Research and Development (UMAGD), 2019. [CrossRef]
5. Europen Assocation of urology. Available online: https://patients.uroweb.org/en/ (accessed on Nowember 2019).
6. Oncology Training and Research Hospital. Available online: http://www.onkoloji.gov.tr (accessed on Nowember 2019).
7. Uro Oncology Society. Available online http://www.uroonkoloji.com/. (accessed on January 2020).
8. Heller, N; Sathianathen, N; Kalapara, A; et al. The KiTS19 Challenge Data: 300 Kidney Tumor Cases with Clinical Context, CT Semantic Segmentations, and Surgical Outcomes, Minnesota University, 2019.
9. Heller, N. Available online: https://github.com/neheller/kits19. (accessed on January 2020).
10. Lamba, H. Understanding Semantic Segmentation with UNET. (accessed on January 2020).
11. Turk, F; Luy, M; Barısci, N. Comparison of Unet and Unet-ResNet Models for Kidney Tumor Segmentation, Confer International Symposium on Multidisciplinary Studies and Innovative Technologieence (ISMSIT) 2019. [CrossRef]
12. Merve, A. K; Bulent, B. A Detailed Analysis on Crowd Analysis with Deep Learning Research [CrossRef]
13. Giovanni, M. Available online: https://github.com/MrGiovanni/UNetPlusPlus. (accessed on January 2020).
14. Vununu, C.; Lee, S.-H.; Kwon, O.-J.; Kwon, K.-R. A Dynamic Learning Method for the Classification of the HEp-2 Cell Images. *Electronics* **2019**, *8*, 850. [CrossRef]
15. Zhou, Z; Rahman, M; et al. Unet++: Redesigning Skip Connections to Exploit Multiscale Features in Image Segmentation.

16. Lee, C.S; Ariel, J; Nicolass, P; Yue, W; Ariel, R; Aaron, Y. L. Deep-learning based, automated segmentation of macular edema in optical coherence tomography [CrossRef]
17. Ravani, K; Saboo, S; Bhatt, J.A Practıcal Approach For Sar Image Despecklıng Usıng Deep Learnıng. IGARSS,2019 Conference, 2019.
18. Smistad, E; Østvik, A; Haugen, B.O; L⊘vstakken, L. 2D left ventricle segmentation using deep learning. Ultrasonics, IEEE Symposium (IUS)