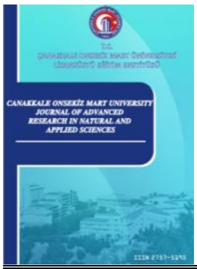
e-ISSN 2757-5195



Çanakkale Onsekiz Mart Univerrsity

Journal of Advanced Research in Natural and Applied Sciences

Open Access

doi.org/10.28979/jarnas.1228740 2023, Vol. 9, Issue 4, Pages: 844-853 dergipark.org.tr/tr/pub/jarnas

Kidney Segmentation with LinkNetB7

Cihan Akyel[1,*]

[1]Ministry of Education, General Directorate of Construction and Real Estate, Ankara, Türkiye

Article History

Received: 03.01.2023

Accepted: 27.07.2023

Published: 20.12.2023

Research Article

Abstract − Cancer is a deadly disease for which early diagnosis is very important. Cancer can occur in many organs and tissues. Renal cell carcinoma (RCC) is the most common and deadly form of kidney cancer. When diagnosing the disease, segmentation of the corresponding organ on the image can help experts make decisions. With artificial intelligence supported decision support systems, experts will be able to achieve faster and more successful results in the diagnosis of kidney cancer. In this sense, segmentation of kidneys on computed tomography images (CT) will contribute to the diagnosis process. Segmentation can be done manually by experts or by methods such as artificial intelligence and image processing. The main advantages of these methods are that they do not involve human error in the diagnostic process and have almost no cost. In studies of kidney segmentation with artificial intelligence, 3d deep learning models are used in the literature. These methods require more training time than 2d models. There are also studies where 2d models are more successful than 3d models in organs that are easier to segment on the image. In this study, the LinkNetB7 model, which has not been previously used in renal segmentation studies, was modified and used. The study achieved a dice coefficient of 97.20%, precision of 97.30%, sensitivity of 97%, and recall of 97%. As a result of the study, LinknetB7 was found to be applicable in kidney segmentation. Although it is a 2d model, it is more successful than UNet3d and some other 2d models.

Keywords – *Decision Support System, image processing, image segmentation, kidney cancer, LinkNetB7.*

1. Introduction

The kidney is a vital organ that filters pollutants from the blood and provides for the excretion of waste through the urine. Although there are several types of kidney cancer, the deadliest and most common form is renal cell carcinoma (RCC). RCC accounts for about 90% of all kidney cancers. Kidney cancer can metastasize to other organs through the bloodstream. For this reason, early diagnosis of kidney cancer is extremely important. It is observed that metastases occur in about 30% of patients when renal cancer is first diagnosed (Demir & Balçık, 2022).

The incidence of kidney cancer may increase due to factors such as smoking and physical inactivity. Early diagnosis of cancer is very important to increase the survival rate of patients and prevent metastasis. Regular screening is important for early detection of kidney cancer. Despite treatment, metastasis can occur in this type of cancer (Kölükçü et al., 2019).

Although the diagnosis can be made by examining CT images, histopathologic examination is required for definitive diagnosis of renal cancer (Devrim, 2019). CT images of the abdominal region are used to detect renal cancer. However, CT images of this region also show other organs such as the liver (Üyetürk et al., 2014). In this case, the differentiation (segmentation) of the kidney on these images becomes important for the diagnostic process.

RCC is the most common urologic cancer with a mortality rate of over 40% (Budak et al., 2013). Diagnosis of RCC can be made by experts by examining ultrasound and/or CT images. Depending on the results, a definitive diagnosis may be made by histopathological examination, if necessary. In the diagnostic process, segmentation

[1] cihan.akyel1@gazi.edu.tr

*Corresponding author

of the kidney in abdominal CT images can be considered as the first step for both classical methods and artificial intelligence-based methods.

To perform kidney segmentation with artificial intelligence algorithms, datasets containing masks suitable for supervised learning are needed. One of these datasets is the dataset named "KiTS19 challenge dataset" in 2019 (KiTS19, 2019). This dataset is preferred in many studies because it is open to the right and allows comparison with other studies.

The biggest challenge for the healthcare industry is to provide quality services at an affordable cost. At this stage, decision support systems (DSS) can help the expert make the right decision. By using artificial intelligence algorithms in decision support systems, greater success can be achieved (Kumar et al. (2011).

There are studies in the literature showing that LinkNet is more successful than UNet in segmenting medical images (Kallam et al. (2020), Akyel and Arıcı (2022). The success of the models can be increased by hybrid use in medical image segmentation. As an alternative to the 2d and 3d models in the existing literature, the LinkNet-based LinkNetB7 model study was used. LinkNetB7 was preferred due to its flexible structure (encoder-decoder architecture), its great success in medical image segmentation, and the fact that it is not used in this field.

1.1. Literature Review

Deep learning models with a seeker-solver architecture provide successful results in medical image segmentation. In models with this structure, features are extracted from images along coding blocks. Solver blocks attempt to obtain output data corresponding to these features. Hsiao et al. (2022) obtained a dice coefficient of 96.90% using the KitS2019 dataset. In the study, an algorithm with an encoder-decoder architecture was developed. EfficientNetB5 was chosen as the encoder model. EfficientNet is an algorithm with high success scores that was introduced by Google in 2019. In the study, images were subjected to contrast enhancement before being imported into the system.

The dataset used for kidney segmentation with Deep Learning can directly influence success. In particular, images where the kidney is not included in the dataset can reduce success. As a solution to this situation, some studies do not include images that do not contain the kidney in the training. In the study by Da Cruz et al. (2021), UNet was the preferred model. In the study, CT images that did not contain the kidney were extracted and not included in the segmentation training. In the study, a dice coefficient of 96.33% was obtained with KiTS2019 data.

In image preprocessing, removing the areas outside the kidney that are classified as noise can have a positive impact on success. Zhao et al. (2020) presented a study on kidney segmentation in 2020. In the study, they used a UNet-based algorithm. It was found that the noise on CT can negatively influence the success. From this point of view, it is found that the noise is cleaned in the data to be trained. In this study, a dice coefficient of 96.69% was obtained using the KiTS19 dataset.

The UNet algorithm can be used as a hybrid with different models. Different models can be used for the coding blocks in the UNet model. This increases the success and thus the number of features that can be extracted from images. A review of the literature shows that models such as UNet and Effi-cient-Net and ResNet can be preferred together. One of the studies that illustrates this type of usage is Li et al. (2022). In the study, residual blocks, which are the basis of ResNet algorithm, are used together with UNet. Using the algorithm called Res-UNet, an accuracy of 96.54% was achieved in the segmentation of kidneys. It was noted that proprietary datasets were used in the study.

Pixel loss in the data being trained is one of the problems encountered in media image segmentation. System constraint can be used to train the images of the dataset by reducing their size. In some studies, the images are divided into slices to solve this situation. In a study by Haghighi et al. (2018), kidney images were segmented. They used a dataset called DCE-MRI, which contains data from 30 patients. One of the most important aspects of the study is the segmentation of each training image. In this study, a dice coefficient of 91.4% was obtained using the UNet model.

LinkNet is a low training time model in the decoder architecture used in image segmentation. There are studies in which it is more successful than the UNet architecture. One of these examples was presented by Akyel and

Arıcı (2022). In their study, about 8% higher detection success was achieved in skin cancer image segmentation than with the standard UNet model. In another study, a modified LinkNet model was used to segment histo-pathological images. According to this study, LinkNet showed about 2% higher success (Kallam et al. 2020).

3d models are generally preferred for segmentation of organs such as kidney and liver. With this default, it is effective that CT images can be trained in slices. However, there are also examples in the literature of 2d models being effective for organs that are easier to segment, such as kidneys. An example of this situation is the work of Zettler and Mastmeyer, (2021). In the study, 2d and 3d UNet models were compared. According to the comparison result, 3d UNet consumes more system memory than 2d UNet. UNet3d, in turn, requires 41 seconds more epoch time. UNet2d achieved about 1% higher success in liver segmentation and 2% higher success in kidney segmentation. The reason for this is the low number of axial slices.

A review of the literature shows that a LinkNet-based algorithm is not used in kidney segmentation. In this study, unlike the literature, the LinkNetB7 model was preferred. In the study using the KiTS19 dataset, the images were divided into 36 patches to avoid pixel loss. Contract enhancement and normalization were applied to the KiTS19 dataset. In the second part of the study, the models and methods used are described in detail. The third section presents the obtained results compared to some studies. The fourth section summarizes the study and discusses what can be done in future studies.

2. Materials and Methods

A publicly available dataset with CT images from KiTS2019 was used for the study. During the training phase, the dataset was split into 80% training and 20% validation. The hyperparameters used are listed in Table 1. The results obtained with the 2d LinkNetB7 model used were combined and converted to a 3d form.

Table 1
Hyper parameters

Parameter	Value
Learning Rate	0.0001
Optimizer	Adam
Output Function	Sigmoid
Loss Function	DLMSE
Epoch Number	50
Input Size	256x256x3

2.1. Dataset

It consists of kidney images CT from 300 patients between 2010 and 2018 in the KiTS19 dataset. These images are in NIFTI format and consist of a varying number of slices. In the study, the images in NIFTI format were converted to jpeg format and used for training. All slices were combined into a single folder without distin-guishing between patients. Information about the KiTS19 dataset that was used can be found in Table 2.

Table 2
KiTS19 Dataset

Parameter	Value
Slice thicknesses range	1mm to 5 mm
Number of files in NIFTI format	300 (210 for Training, 90 for Test)
Number of images in jpeg format	432000 (360000 for Training, 72000 for Test)
Licence information	Public CC BY-NC-SA

2.2. Pre-Processing Phase

First, normalization or contrast enhancement was performed with Clahe to distinguish the kidney from the abdominal cavity by increasing the contrast difference between the images. This process is shown in Fig 1.

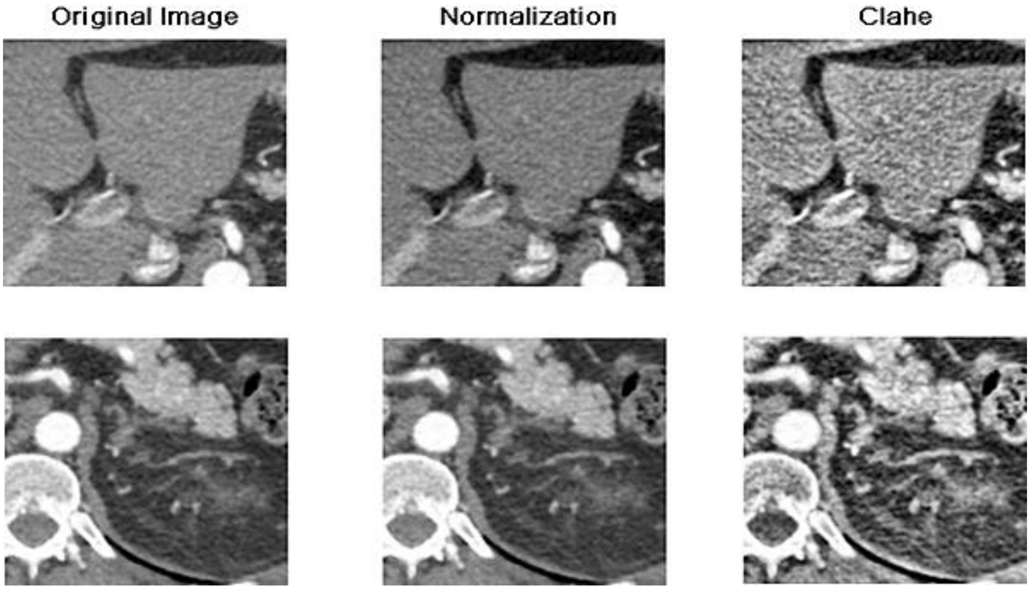


Figure 1. Image processing steps

In the next step, the images of the dataset are divided into 36 layers with a resolution of 256x256. With this method, there is no pixel loss for images up to 1536x1536 resolution, and pixel loss is reduced for higher resolutions. In addition, the images that could not contain kidneys in the training images of the study were not included in the system according to the mask value with the image processing technique. This process is shown in Fig 2.

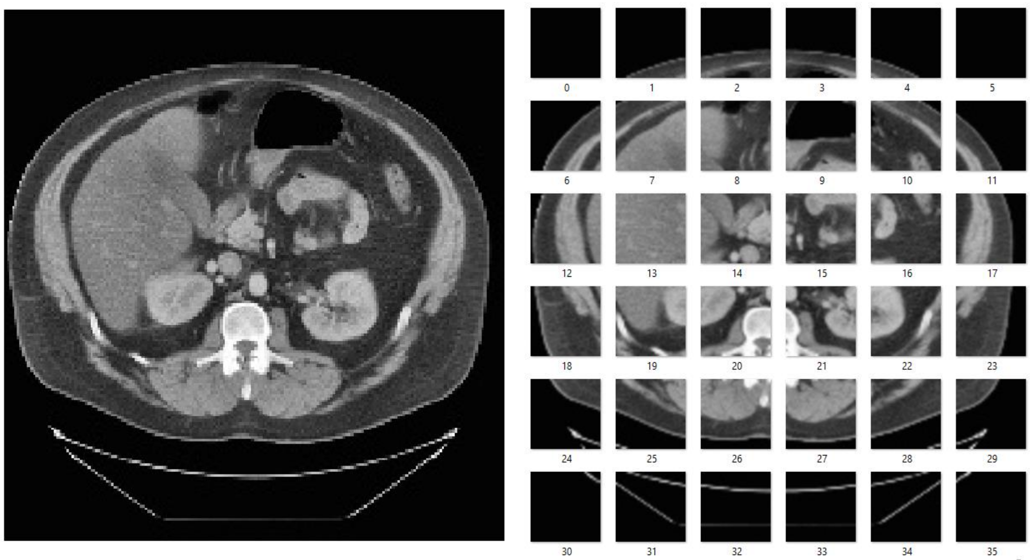


Figure 2. The process of dividing the image into pathces

2.3. LinkNet and LinkNetB7 Architecture

Two features that distinguish LinkNet are its low training time and its success in medical image segmentation. This model was proposed in 2017. It consists of 4 coding blocks and 4 decoding blocks. In the coding blocks, Res-Net18 is preferred to reduce the epoch time. As with the UNet model, LinkNet can also achieve an increase in success by choosing different coding models (Chaurasia & Culurciello, 2017). The LinkNet architecture is shown in Fig 3.

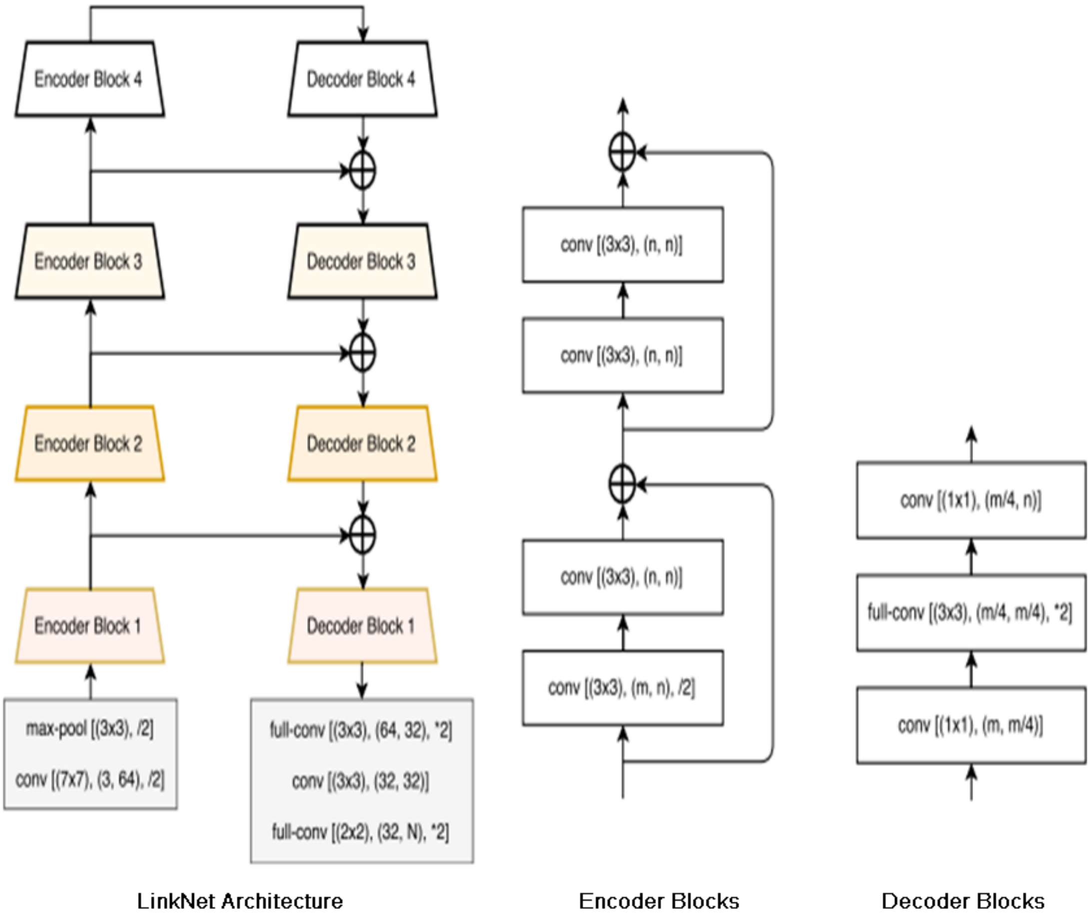


Figure 3. LinkNet architecture (Chaurasia & Culurciello, 2017)

The LinkNetB7 algorithm is a LinkNet-based model. In this model, EfficientNetB7 is preferred as the encoder. With the encoder blocks, the input image is reduced to half its size by pooling at each station. The attributes are extracted from the images in the encoder blocks. The number of attributes extracted with EfficientNetB7 has been increased. In terms of cost, the epoch time was higher than for the standard LinkNet.

A middle block was added to the architecture, which was not included in the original model. The middle block consists of a total of 3 3x3 convolutional layers with 1024 cores. Feature extraction is also performed between the middle block and the encoder and decoder blocks. The image downsampled to 8x8 resolution in the decoder blocks reaches 2x resolution with upsampling at each station. The activation function Relu is used.

Although the ResNet block added before the last layer increases the success by 0.1%, it was removed from the original LinkNetB7 architecture because it increases the epoch time and makes the algorithm cumbersome. Figure 4 shows the architecture of the algorithm. Encoder layers can be seen in Table 3.

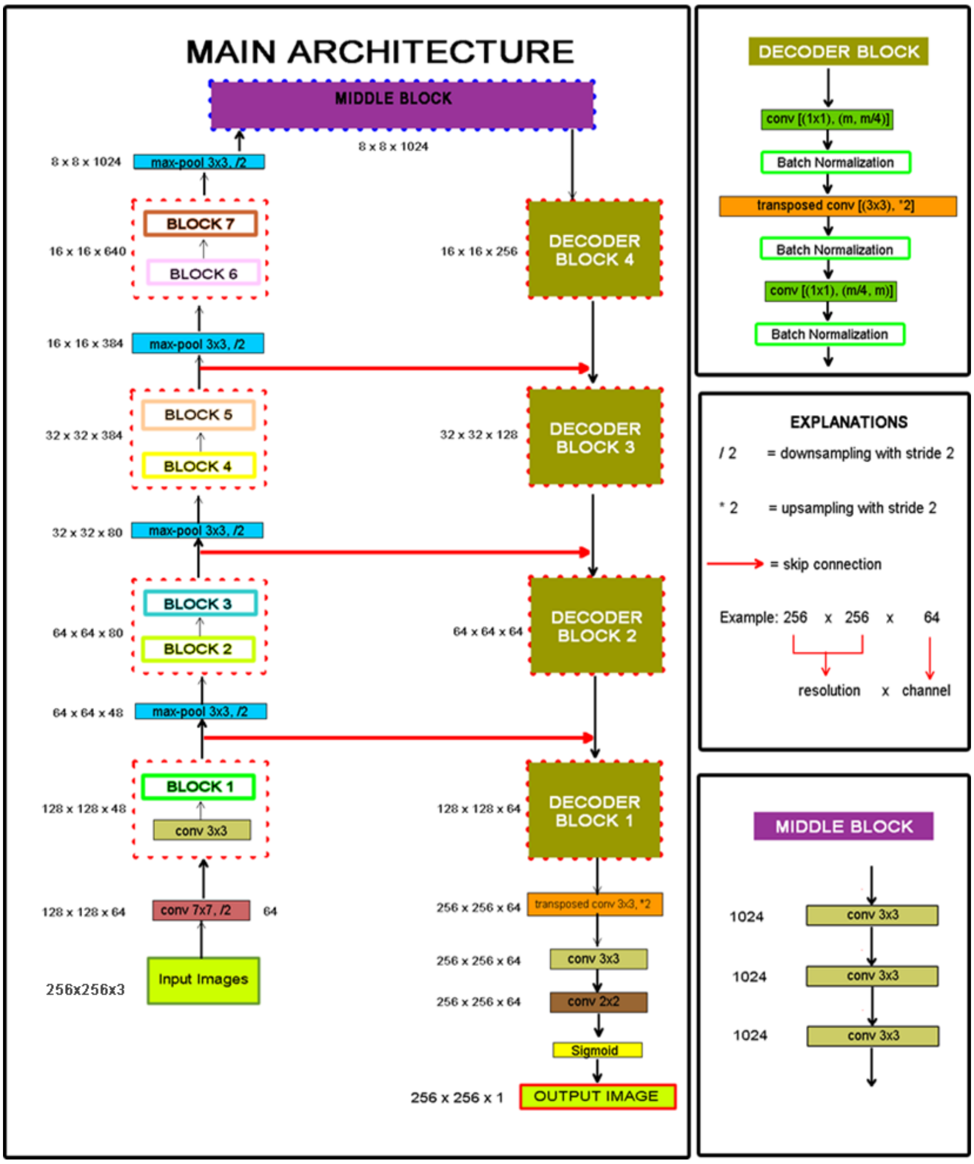


Figure 4. LinkNetB7 architecture (Akyel, & Arıcı, 2022)

Table 3
Encoder Layers

Phase	Operator	Resolution	Channels	Layers
1	Conv 7x7,/2	128 x 128	64	1
2	Conv 3x3	128 x 128	64	1
3	Block 1 – MBconv1 3x3	128 x 128	32	3
4	Block 2 – MBconv6 3x3	64 x 64	48	7
5	Block 3 – Mbconv6 5x5	64 x 64	80	7
6	Block 4 – Mbconv6 3x3	32 x 32	80	10
7	Block 5 – Mbconv6 5x5	32 x 32	224	10
8	Block 6 – Mbconv6 5x5	16 x 16	384	13
9	Block 7 – Mbconv6 3x3	16 x 16	640	4

2.4. Metrics

2.4.1. Dice Coefficient

Dice coefficient calculates the match between the real (Y) and predicted (X) segmentation areas. The formula for dice coefficient is given in Equation 2.1 (Akyel, & Arıcı, 2022).

$$\text{Dice Coefficient} = (2 * |X \cap Y|)/(|X| + |Y|) \quad (2.1)$$

2.4.2. Loss Function

In this study, instead of using a single loss function, two different functions are combined and their advantages are combined. The loss functions used are dice loss (DL) and mean square error (MSE). The main advantage of die loss is that it is successful against oversegmentation errors. MSE, on the other hand, can increase success by reducing overall image details. As can be seen in Equation 2.2, the two functions are combined under the name DLMSE (Akyel, & Arıcı, 2022).

$$DLMSE = MSE + DL \tag{2.2}$$

2.4.3. Optimizer

In the study, adam was chosen as the optimizer. Adam is a stochastic gradient descent method (Akyel, & Arıcı, 2022). Adam can achieve high success from less epoch number.

2.4.4. Recall

Determines the accuracy rate of positively predicted data in the test dataset being actually positive (Akyel, & Arıcı, 2022).

$$\text{Recall} = \text{True Positive}/(\text{True Positive} + \text{False Negative}) \tag{2.3}$$

2.4.5. Precision

It is a measure of how positive predicted values in the test data set actually are. In Equation 2.4, the equation for calculating the estimation value is given (Akyel, & Arıcı, 2022).

$$\text{Precision} = \text{True Positive}/(\text{True Positive} + \text{False Positive}) \tag{2.4}$$

3. Results and Discussion

The model was run for 50 epochs with the parameters given in Table 1. Table 4 shows the results obtained in comparison with some other studies. And in Figure 5, training graphic can be seen. For the segmentation of images composed of layers, 3d models come to the fore because of their compatibility with this structure. 2d models, on the other hand, can be used for segmentation of CT images by considering all the sliced images as a single dataset. The biggest advantage of 2d models over 3d models is that they require less system resources. There are examples such as Zettler and Mastmeyer, (2021) where the 2d model is more successful than the 3d model in segmenting CT images of the liver. A look at Table 4 shows that the model used has a high degree of success. The main reason why the used model is more successful is that the blocks belonging to EfficientNetB7 were selected as encryption blocks.

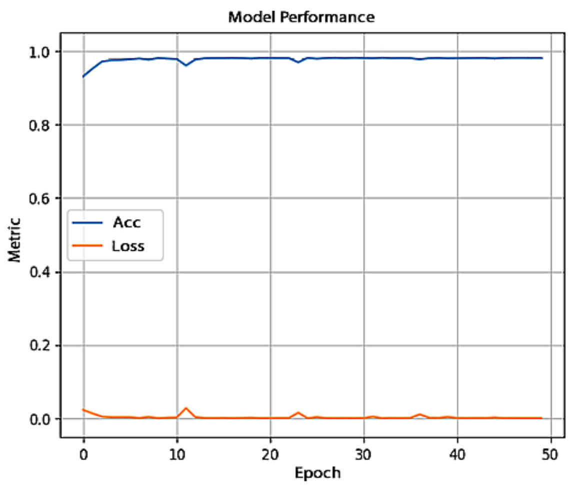


Figure 5. Training graphic

Table 4
Comparison Results

Reference	Method	Dice Coefficient (%)	Precision (%)	Sensitivity (%)	Recall (%)	Dataset
Hsiao et al., 2022	EfficientNetB5	96.90	97.47	-	96.45	KiTS19
Da cruz et al.., 2020	UNet2d	96.33	-	95.32	-	KiTS19
Zhao et al., 2020	UNet3d	96.90	97.10	-	96.80	KiTS19
Li et al., 2022	ResUnet	96.54	-	96.49	-	Own
Haghighi et al., 2018	UNet3d	87.50	92.7	-	-	DCE-MRI
UNet2d	UNet2d	96,50	96,55	95,90	96.20	KiTS19
UNet3d	UNet3d	96,80	96,85	96,10	96.25	KiTS19
Used Model	LinkNet	96.62	96.58	96.97	96.18	KiTS19
Used Model	**LinkNetB7**	**97.20**	**97.30**	**97**	**97**	**KiTS19**

3.1. Ablation Study

In this study, an ablation study was performed to investigate the effects of segmentation (ES), use of EfficientNetB7 as an encoder (EB7), contrast enhancement, and normalization methods (CN) on success. The results are shown in Table 5. Figure 6 shows examples of mask estimates.

Table 5

Ablation Study

Structure	Dice Coefficient (%)
LinkNetB7	90
LinkNetB7 + ES	91.75
LinkNetB7 + EB7	94.05
LinkNetB7 + KN	91.80
LinkNetB7 + ES	92.04
LinkNetB7 + ES + EB7	95,35
LinkNetB7 + ES + EB7 + CN	**97.20**

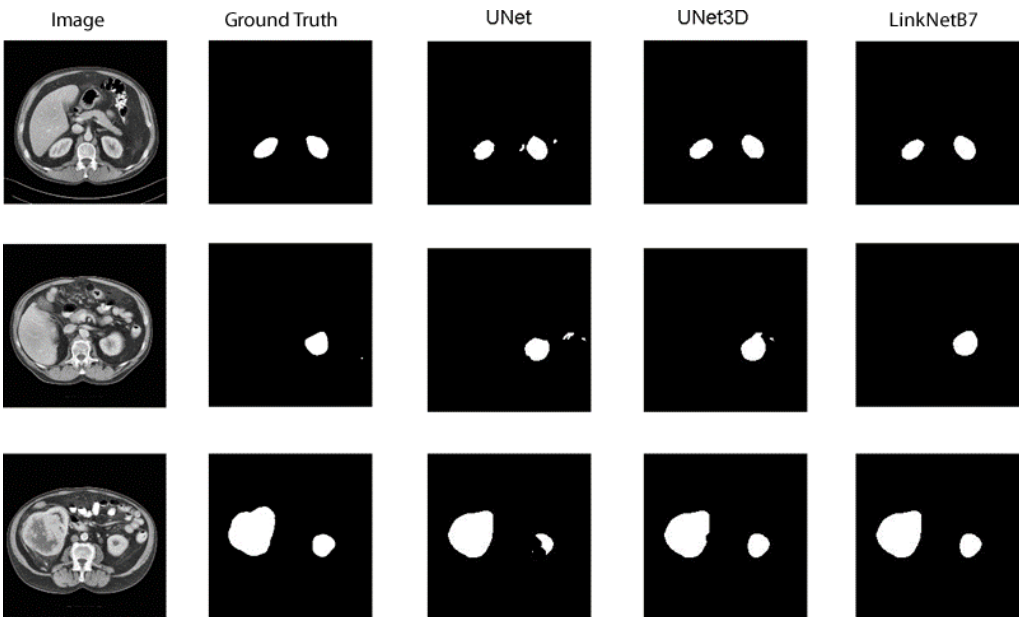


Figure 6. Sample segmentation results

4. Conclusion

The results obtained in the study show that the LinkNetB7 model has a higher success than the other compared models. The success was 0.4% higher than that of the 3d-UNet model. The main reason is that segmentation is easier in kidney images and the number of layers is less. In addition, the success of the architecture in segmentation was also higher compared to UNet3d. There is no other study in the literature that uses LinkNetB7 for kidney segmentation. Therefore, this study found that the LinkNetB7 model was suitable for kidney segmentation. LinkNetB7 that is 2d model was preferred because of its accuracy in medical segmentation. Also 2d model requires fewer system memory than 3d models.

As can be seen in Table 6, the 2d LinkNetB7 model used requires less epoch time than 3d UNet + EfficientNetB7. Although the basic UNet model requires relatively less epoch time than LinkNet, the fact that LinkNet is more successful than UNet puts the LinkNet model forward. Also it was seen that the LinkNetB7 model has high success values in kidney segmentation.

Table 6
Training times

Model	Epoch Time (min)
UNet3d	9.60
UNet3d + EfficientNetB7	13.10
LinkNet	8.70
UNet2d	8.55
LinkNetB7	11.10

As the resolution and number of images in the image datasets used in deep learning methods increase, the systems become insufficient. The slicing method can be used to solve this situation. In future studies, the system can be trained by using more epochs with higher input size in a system with better hardware. In future studies, different encoder models can be tried to consume less resources without decreasing the success rate.

Author Contributions

Cihan Akyel: prepared the data, created the model, analyzed the results and wrote the paper.

Conflicts of Interest

The authors declare no conflict of interest.

References

Akyel C, & Arıcı N. (2022). LinkNet-B7: Noise Removal and Lesion Segmentation in Images of Skin Cancer. *Mathematics, 10*(5), 736-751. DOI: https://doi.org/10.3390/math10050736

Budak, S., Sağlam, H. S., Köse, O., Kumsar, Ş., & Adsan, Ö. (2013). Böbrek Hücreli Karsinomada Radyolojik Tümör Boyutu ile Patolojik Boyutun İlişkisi, *Sakarya Medical Journal, 3(*4), 186-189, DOI: https://doi.org/10.5505/sakaryamj.2013.92485

Chaurasia, A., & Culurciello, E. (2017). LinkNet: Exploiting The encoder Representations for Efficient Semantic Segmentation. T*he IEEE Visual Communications and Image Processing (VCIP).* St. Petersburg, FL, USA, Retrieved from: https://arxiv.org/abs/1707.03718

Da Cruz, L. B., Araújo, J. D. L., Ferreira, J. L., Diniz, J. O. B., Silva, A. C., De Almeida, J. D. S., De Paiva, A. C., & Gattass, M. (2020). Kidney segmentation from computed tomography images using deep neural network, *Computers in Biology and Medicine, 123,* 1-12, DOI: https://doi.org/10.1016/j.compbiomed.2020.103906.

Demir, H., & Balçık, P. Y. (2022). Ödeyici Kurum Bakış Açısıyla İleri Evre Böbrek Kanserinde Kaynak Kullanımı: Eğitim and Araştırma Hastanesi Örneği, *Vizyoner Journal, 13*(34), 1-15. DOI: https://doi.org/10.21076/vizyoner.991598

Devrim, T. (2019). Böbrek Tümörü Vakalarının Retrospektif Olarak Değerlendirilmesi, *Kırıkkale Üniversitesi Tıp Fakültesi Journal, 21*(2), 212-217, DOI: https://doi.org/10.24938/kutfd.552211

Haghighi, M., Warfield, S. K., & Kurugol, S. (2018). Automatic renal segmentation in DCE-MRI using convolutional neural networks, *2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018* (pp. 1-9). Washington, USA. Retrieved from: https://ieeexplore.ieee.org/document/8363865

Hsiao, C., Lin, P., Chung, L., Lin, F. Y., Yang, F., Yang, S., Wu, C., Huang, Y., & Sun, T. (2022). A deep learning-based precision and automatic kidney segmentation system using efficient feature pyramid networks in computed tomography images, *Computer Methods and Programs in Biomedicine, 221*, 1-17, DOI: https://doi.org/10.1016/j.cmpb.2022.106854

Kallam, S., Kumar, M.S., Natarajan, V.A., & Patan, R. (2020). Segmentation of Nuclei in Histopathology images using Fully Convolutional Deep Neural Architecture. *The 2020 International Conference on Computing and Information Technology (ICCIT-1441)* (pp. 319-325). Tabuk, Saudi Arabia. Retrieved from: https://ieeexplore.ieee.org/document/9213817

Kumar, S., Sathyadevi, G. & Sivanesh, S. (2011). "Decision Support System for Medical Diagnosis Using Data Mining," Retrieved from: https://www.researchgate.net/publication/267934132

Kölükçü, E., Deresoy, F. A., Uluocak, N., Atılgan, D., Gümüşay, Ö., Beyhan, M., & Kılıç, Ş. (2019). Sarkomatoid renal hücreli karsinom: olgu sunumu and literatürün gözden geçirilmesi, *Anadolu Güncel Tıp Journal*, *1*(2), 42-46, DOI: https://doi.org/10.38053/agtd.512103

Li, D., Xiao, C., Liu, Y., Chen, Z., Hassan, H., Su, L., Li, H., Xie, W., Zhong, W., & Huanglund, B. (2022). Deep Segmentation Networks for Segmenting Kidneys and Detecting Kidney Stones in Unenhanced Abdominal CT Images, *diagnostics*, *12*, 1-17, DOI: https://doi.org/10.3390/diagnostics12081788

The KiTS19 Challenge Data: 300 Kidney Tumor Cases with Clinical Context, CT Semantic Segmentations, and Surgical Outcomes. Retrieved from: https://arxiv.org/abs/1904.00445

Üyetürk, Ü., Üyetürk, U., & Metin, A. (2014). Böbrek Hücreli Kanser, *Gaziosmanpaşa Üniversitesi Tıp Fakültesi Journal, 6(*1), 1-17. Retrieved from: https://dergipark.org.tr/tr/pub/gutfd/issue/34239/378385

Zhao, W., Jiang, D., Queralta, J. P., & Westerlund, T. (2020). MSS U-Net: 3d segmentation of kidneys and tumors from CT images with a multi-scale supervised U-Net, *Informatics in Medicine Unlocked, 19,* 1-11, DOI: https://doi.org/10.1016/j.imu.2020.100357

Zettler, N. & Mastmeyer, A. (2021). Comparison of 2d vs. 3d U-Net Organ Segmentation in abdominal 3d CT images. Retrieved from: arxiv.org/abs/2107.04062