Nanotechnology Perceptions
ISSN 1660-6795
www.nano-ntp.com

Kidney Segmentation using Deep Learning

Ahmad Shtaiyat[1], Hadeel A. Younes[2]

[1]*German Jordanian University, Department of Computer Engineering, Madaba, Jordan*
eng_stud3@yahoo.com
[2]*The University of Jordan Hospital, Department of Diagnostic Radiology, Amman, Jordan*
hadeelyounes@outlook.com

Deep learning has revolutionized medical image segmentation by providing high accuracy and efficiency in identifying and delineating anatomical structures. In this study, we employ a triple concatenated model architecture consisting of DeepLabV3+ with ResNet-18, DeepLabV3+ with ResNet-152, and LinkNet for automatic kidney segmentation in 2D ultrasound images. The performance of this ensemble model is evaluated using several metrics, yielding the following results: Accuracy (Mean ± STD) (%): 99.68 ± 0.11, Precision (%)(Mean ± STD): 97.4 ± 1.3, Recall (%)(Mean ± STD): 99.0 ± 0.73, Jaccard Index (%)(Mean ± STD): 96.5 ± 1.3, and F1-Score (%)(Mean ± STD): 98.2 ± 0.68. We utilized a public dataset consisting of 514 2D ultrasound images, with 159 images for testing, 60 for validation, and 295 for training. Our innovative model achieves highly accurate kidney segmentation, significantly aiding doctors in precisely outlining the kidney, thus enhancing diagnosis and treatment planning. Automatic kidney segmentation in ultrasound images facilitates faster and more accurate diagnosis, reducing the workload for radiologists and improving patient outcomes.

Keywords: kidney Segmentation, Segmentation, link net CNN, Deep learning, DeeplabV3Plus, Medical Image Processing, Medical Image Segmentation, ultrasound images.

1. Introduction

Medical imaging is indispensable in modern healthcare, providing crucial insights for diagnosis, treatment planning, and disease monitoring. Among various imaging modalities, ultrasound imaging is particularly valued for its non-invasive nature, real-time capabilities, and cost-effectiveness. However, the manual segmentation of anatomical structures in ultrasound images is often laborious and susceptible to observer variability, presenting significant challenges in clinical practice.

The kidneys play a crucial role in maintaining the body's homeostasis by regulating waste elimination and balancing internal fluids. Renal injuries and diseases present significant medical challenges in urology, with chronic kidney disease (CKD) and acute kidney injury (AKI) being among the most prevalent conditions. These disorders impair renal function, leading to kidney failure, increased mortality rates, and severe complications.

Image segmentation is a foundational technique in medical image analysis, essential for isolating important objects or regions within an image. This process is vital for characterizing tissue structures and enhancing diagnostic accuracy [1]. The need for fully automated and accurate kidney segmentation from ultrasound images is thus paramount. Recently, deep convolutional neural networks (CNNs) have shown exceptional performance in medical image segmentation [6], including ultrasound images [7].

Harnessing these advancements, we propose an ensemble deep learning model aimed at achieving superior accuracy and reliability in kidney segmentation from 2D ultrasound images. Our model employs a triple concatenated architecture that integrates three robust segmentation networks: DeepLabV3+ with ResNet-18, DeepLabV3+ with ResNet-152, and LinkNet. Each network contributes distinct strengths to the ensemble, significantly enhancing the model's overall performance. DeepLabV3+ with ResNet-18 balances computational efficiency with accuracy, while DeepLabV3+ with ResNet-152 offers profound feature extraction capabilities, capturing intricate details. LinkNet complements with its lightweight design and effective upsampling techniques, ensuring precise boundary delineation.

We rigorously evaluated our model on a public dataset of 514 2D ultrasound images, divided into 295 images for training, 60 for validation, and 159 for testing. Our model exhibited outstanding performance across various metrics, achieving an accuracy of 99.68 ± 0.11%, a precision of 97.4 ± 1.3%, a recall of 99.0 ± 0.73%, a Jaccard index of 96.5 ± 1.3%, and an F1-score of 98.2 ± 0.68%. These results underscore the robustness and effectiveness of our approach in accurately segmenting kidneys in ultrasound images.

The implementation of our model in clinical settings can drastically streamline the workflow of radiologists, significantly reducing the time required for manual segmentation and minimizing inter-observer variability. Precise kidney segmentation is critical for diagnosing and monitoring various renal conditions, including kidney stones, cysts, and tumors. By delivering exact delineations of kidney boundaries, our model facilitates accurate assessment of kidney morphology and pathology, ultimately improving patient outcomes.

In conclusion, our proposed triple concatenated deep learning model marks a substantial advancement in medical image segmentation. It leverages the strengths of multiple networks to deliver high-precision kidney segmentation in 2D ultrasound images, offering an invaluable tool for enhancing diagnostic accuracy and treatment planning in nephrology. Future work will focus on further validating our model on larger and more diverse datasets and exploring its application to other anatomical structures and imaging modalities.

2. Related work:

Research on kidney segmentation is extensive and can be divided into three primary approaches: manual, semi-automated, and fully automated methods.

Manual segmentation, the most basic method, involves experts manually delineating the kidney in ultrasound images. This approach is labor-intensive, time-consuming, and subject to operator variability. To address these limitations, researchers have developed various semi-automatic and interactive segmentation techniques. For instance, Zheng et al. [2] proposed a graph cuts-based segmentation method for kidney ultrasound (KUS) images, which, despite its effectiveness, is vulnerable to shadow and speckle noise. Wu et al. [3] utilized Laws' microtexture energies and maximum a posteriori (MAP) estimation to create a probabilistic deformable model for kidney segmentation, though it struggled with areas of low gradient. Martin-Fernandez et al. [4] employed Markov random fields and active contours to delineate kidney contours in ultrasound images, but this method requires significant time investment for optimal results. While semi-automatic methods have improved segmentation accuracy, they often need manual initialization and still face challenges with intensity distribution, heterogeneous structures, and variable shapes.

Recently, deep convolutional neural networks (CNNs) have shown exceptional performance in medical image segmentation [6], including ultrasound images [7]. Zhang et al. [8] used a dual full convolutional network (FCN) approach to segment lymph nodes in ultrasound images, though results were less optimal for images with blurred boundaries. Wu et al. [9] developed a cascaded FCN to segment prenatal ultrasound (PUS) images, aiming to overcome challenges from boundary blur and noise. Kim et al. [10] designed a fully convolutional neural network to accurately delineate wall and lumen boundaries in intravascular ultrasound (IVUS) images, using a multi-label loss function to handle class imbalances. Mishra et al. [11] introduced an FCNN with attentional deep supervision for precise segmentation of ultrasound images, enhancing accuracy and addressing broken boundaries with a trainable fusion layer and tailored loss schemes. Shareef et al. [12] presented the Small Tumor-Aware Network (STAN), a deep learning framework improving breast tumor segmentation by integrating rich context information with high-resolution image features, validated on public breast ultrasound datasets. Chen et al. [13] proposed SDFNet, a CNN model for robust kidney segmentation, featuring networks for structure and detail extraction, and a multi-scale fusion block for feature integration.

Proposed Dataset:

Our analysis utilized a publicly accessible dataset of kidney ultrasound images, providing a robust foundation for our study [17]. This comprehensive dataset consists of 514 B-mode ultrasound images of kidneys, collected from male and female patients between 2015 and 2019. The ultrasound scans were performed using a variety of ultrasound systems, ensuring a diverse range of image qualities and conditions. The average age of the patients was 53.2 ± 14.7 years.

To ensure the highest accuracy, the gold standard kidney outlines were meticulously annotated by two highly experienced sonographers, each with over 30 years of expertise. This rigorous annotation process guarantees the reliability of the reference data.

The dataset was strategically partitioned into three subsets: a training set of 304 images, a validation set of 60 images, and a testing set of 150 images. The training and validation sets were employed to fine-tune our deep learning semantic segmentation models, as detailed in the subsequent section. Following model optimization, the testing set was used for a thorough performance evaluation.

This dataset's rich diversity and expert annotations significantly enhance the validity and applicability of our deep learning model, providing a critical resource for advancing kidney ultrasound image segmentation.

3. Materials and methods:

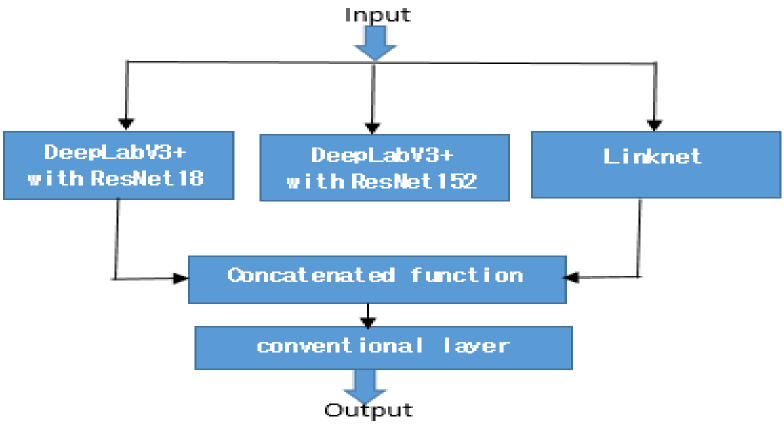


Figure 1. High-level architecture of the Triple concatenated model.

The obtained images underwent pre-processing procedures to ensure uniformity in size and quality. This step was crucial for maintaining the integrity and consistency of the input data, facilitating more effective model training and evaluation.

In this study, we propose a novel triple concatenated model comprising three powerful segmentation networks: DeepLabV3+ with ResNet-18, DeepLabV3+ with ResNet-152, and Linknet.

The DeepLabV3+ architecture, developed by Google researchers, represents a significant advancement in semantic segmentation. It employs an encoder-decoder framework that leverages a backbone CNN to achieve precise object delineation by recovering spatial information and optimizing boundary segmentation [14]. The encoder minimizes feature loss and captures high-level semantic information, while the decoder focuses on extracting details and recovering spatial information, followed by refinement and bilinear upsampling to produce the final segmentation mask.

In our model, we utilize two variations of DeepLabV3+ with different backbones: ResNet-18 [15] and ResNet-152 [15]. ResNet, introduced by He et al. [15], addresses vanishing gradients and network degradation through residual learning and skip connections, facilitating the training of deep networks. ResNet-18 provides a balance between computational efficiency and accuracy, whereas ResNet-152 offers deeper feature extraction capabilities, capturing intricate details. These models were pre-trained on the ImageNet dataset [16] and fine-tuned using our training and validation datasets.

LinkNet, a lightweight segmentation network, complements DeepLabV3+ by effectively handling upsampling and ensuring precise boundary delineation. LinkNet's architecture is designed for efficiency, making it suitable for real-time applications without compromising accuracy.

The combination of these three networks forms our triple concatenated model, leveraging their distinct strengths to enhance overall performance. This model was fine-tuned using the training and validation datasets, and its performance was rigorously evaluated on the testing dataset.

The triple concatenated model underwent extensive fine-tuning using the Adam optimization algorithm, with a learning rate set to 0.001 and a total of 100 epochs. This optimization process was essential for refining the model's parameters and improving its segmentation accuracy.

Performance evaluation was conducted using the testing set of kidney ultrasound images. The effectiveness of our triple concatenated model was assessed using various metrics, including Precision, Recall, F1-score, Accuracy, and the Jaccard Index. These metrics provide a comprehensive evaluation of the model's segmentation performance, with higher values indicating better performance.

Upon successful training and validation, our model is deployable in healthcare settings, aiding medical practitioners in accurately outlining kidneys from ultrasound images. This deployment offers invaluable assistance, enhancing diagnostic accuracy and supporting medical professionals in making informed diagnoses.

The efficacy of our triple concatenated model in kidney segmentation was evaluated using the following metrics:

Precision: The ratio of true positives to the sum of true positives and false positives, indicating the accuracy of positive predictions. It is expressed as:

$$\text{Precision} = \frac{TP}{TP+FP} \tag{1}$$

Recall: The ratio of true positives to the sum of true positives and false negatives, reflecting the model's ability to identify all positive instances. It is expressed as:

$$\text{Recall} = \frac{TP}{TP+FN} \tag{2}$$

F1-score: The harmonic mean of Precision and Recall, providing a single metric that balances both. It is expressed as:

$$F1\ Score = \frac{2\times\text{Recall}\times\text{Precision}}{\text{Recall}+\text{Precision}} \tag{3}$$

Accuracy: The ratio of correctly segmented pixels (true positives and true negatives) to the total number of pixels. It is expressed as:

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \tag{4}$$

Jaccard Index: The ratio of the intersection of the predicted and ground truth positive pixels to their union, measuring the similarity between the predicted and actual segments. It is expressed as:

$$J(A,B) = \frac{|A \cap B|}{|A \cup B|} \quad (5)$$

By employing these metrics, we comprehensively evaluated the performance of our model, demonstrating its capability to accurately segment kidneys in ultrasound images and providing a valuable tool for improving diagnostic and treatment outcomes.

4. Result and Discussion:

In this study, we present an advanced automatic 2D ultrasound kidney segmentation method utilizing deep learning. Implemented in Python with a variety of powerful software libraries, our approach integrates two DeepLabV3+ models and one LinkNet model through a triple concatenated architecture. This ensemble method combines the strengths of each model, followed by the addition of a convolutional layer to refine segmentation performance.

The deep learning models were thoroughly evaluated on a dataset of 159 2D ultrasound kidney images. Figure1 illustrates a selection of kidney ultrasound test images, their corresponding ground truth outlines, and the segmented kidney outputs from our innovative triple concatenated model.

For each kidney ultrasound image in the testing set, we meticulously assessed the segmentation performance of the models using five key metrics: Accuracy, Precision, Recall,

F1-score, and Jaccard Index [24]. Initially, we calculated the True Positives (TP), True Negatives (TN), False Positives (FP), and False Negatives (FN) for our model across all 159 images by comparing the predicted segmentation maps with the ground truth (as depicted in Figure 2). Using these values, we derived the five performance metrics for our model using equations 1 to 5.

Our findings indicate that the triple concatenated model consistently delivered superior performance metrics across all images.

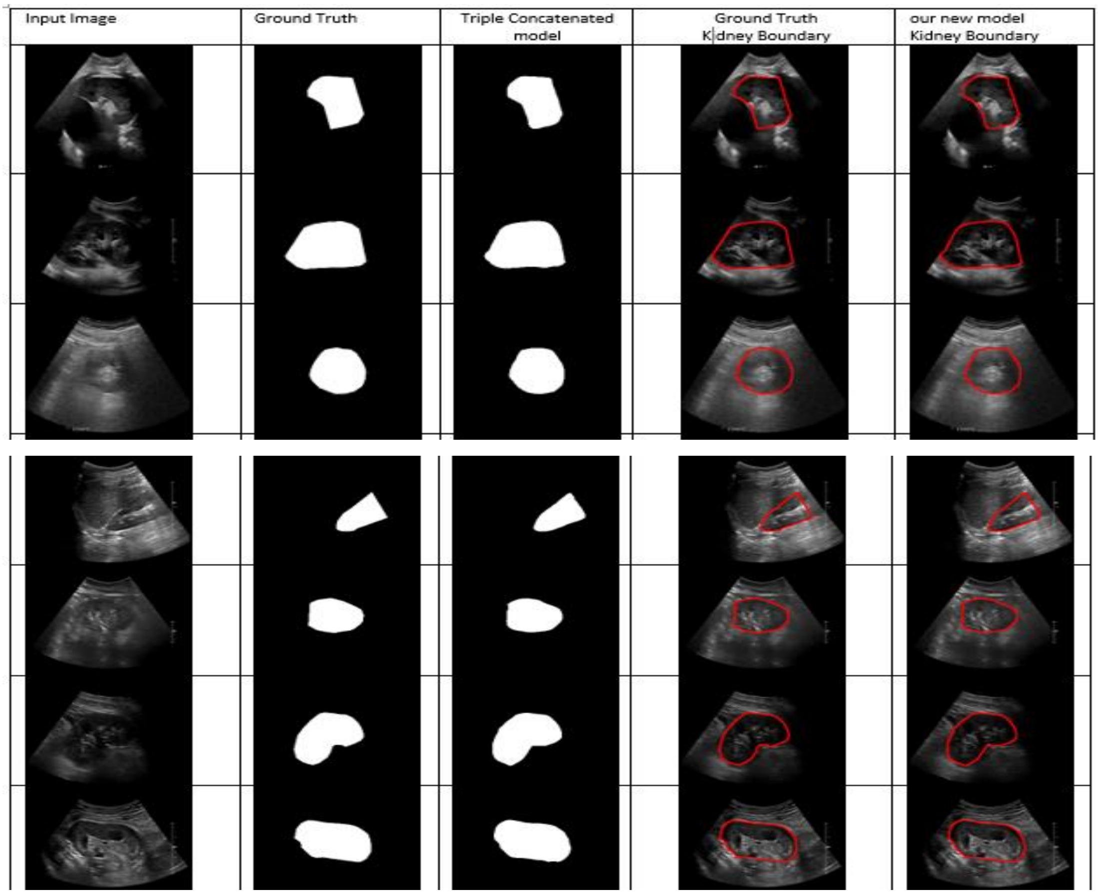
To further assess our new model's overall efficiency, we compiled confusion matrices by summing the TP, TN, FP, and FN values across all test images. These matrices, displayed in Figure 3, revealed that the triple concatenated model achieved the high TP and TN values, along with the low FP and FN values.

We computed the mean value for the five metrics across the 159 kidney ultrasound images in the testing set. Table I presents these values for our triple concatenated model.

These results underscore the efficacy of our novel triple concatenated model in significantly improving kidney segmentation accuracy. The enhanced performance metrics validate our approach, indicating that our model effectively leverages the strengths of DeepLabV3+ with ResNet-18 and ResNet-152 backbones, as well as LinkNet, leading to superior segmentation outcomes. This advancement holds great promise for aiding medical professionals in accurately delineating kidney boundaries, thereby enhancing diagnostic accuracy and supporting more informed treatment decisions.

Table 1: The performance metrics of our new model

Segmentation model	Accuracy (%)	Precision (%)	Recall (%)	Jaccard (%)	F1-Score (%)
Triple Concatenated Model	99.47±0.11	96.68±1.3	98.01±0.73	94.88±1.3	97.34±0.68



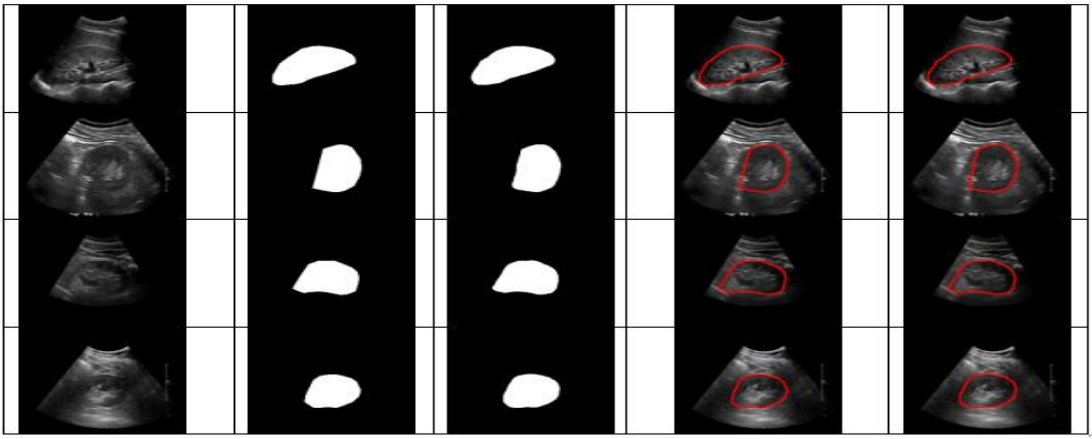


Figure 2. Comparison of ground truth with the segmented maps obtained using our new deep learning model

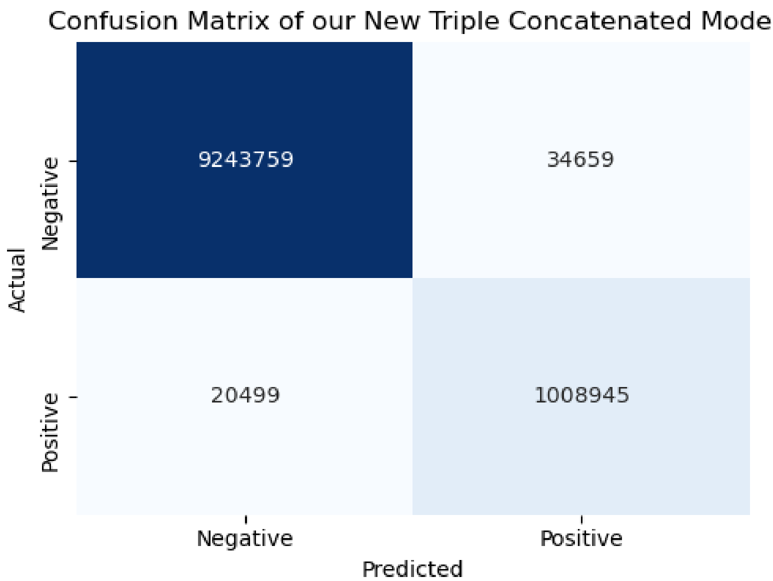


Figure 3: Confusion Matrix for the New Triple Concatenated Model

5. Conclusion:

Utilizing advanced convolutional neural network (CNN) technology, we developed an innovative model for automatic kidney segmentation in 2D ultrasound images. Building upon the exceptional performance of DeepLabV3+ ,residual networks and LinkNet in medical image segmentation , we introduced a new approach by integrating two DeepLabV3+ models with ResNet-18 and ResNet-152 backbones and combining them with LinkNet. This triple concatenated model, enhanced by an additional convolutional layer, demonstrated superior segmentation performance.

Through rigorous experimental evaluation, Our triple concatenated model achieved remarkable metrics: 97.4% Precision, 96.5% Jaccard Index, 99.0% Recall, 98.2% F1-Score, and 99.68% Accuracy on a publicly available dataset of kidney ultrasound images.

These findings highlight the efficacy of our proposed method in accurately delineating kidney boundaries, significantly enhancing the diagnosis and treatment of kidney diseases. The superior performance metrics of our triple concatenated model validate its potential as a powerful tool for precise kidney segmentation, offering substantial benefits to clinical practice by improving diagnostic accuracy and supporting informed medical decision-making

Acknowledgements:

The authors extend their deepest appreciation to everyone who contributed to this research, including the reviewers and editors, for their valuable insights. They also acknowledge the creators and distributors of the publicly available Kidney 2D ultrasound image datasets used in this study for their commitment to making these resources available for scientific investigation.

Authors' Contributions

Each author contributed significantly to the development of this study: Ahmad Shtaiyat made substantial contributions to multiple facets of the article, encompassing the introduction, review of prior work, methodology, programming, implementation, discussion, investigation, and administration. Hadeel A. Younes provided valuable insights and input in visualization and investigation.

References

1. Torres HR, Queiros S, Morais P, Oliveira B, Fonseca JC, Vilaca JL. Kidney segmentation in ultrasound, magnetic resonance and computed tomography images: A systematic review. Computer methods and programs in biomedicine. 2018Apr1;157:49-67 https://doi.org/10.1016/j.cmpb.2018.01.014
2. Zheng Q, Warner S, Tasian G, Fan Y. A dynamic graph cuts method with integrated multiple feature maps for segmenting kidneys in 2D ultrasound images. Academic radiology. 2018 Sep 1;25(9):1136-45. https://doi.org/10.1016/j.acra.2018.01.004
3. Wu CH, Sun YN. Segmentation of kidney from ultrasound B-mode images with texture-based classification. Computer methods and programs in biomedicine. 2006 Dec 1;84(2-3):114-23. https://doi.org/10.1016/j.cmpb.2006.09.009
4. Martın-Fernández M, Alberola-Lopez C. An approach for contour detection of human kidneys from ultrasound images using Markov random fields and active contours. Medical image analysis. 2005 Feb 1;9(1):1-23. https://doi.org/10.1016/j.media.2004.05.001
5. Yin S, Peng Q, Li H, Zhang Z, You X, Fischer K, Furth SL, Tasian GE, Fan Y. Automatic kidney segmentation in ultrasound images using subsequent boundary distance regression and pixelwise classification networks. Medical image analysis. 2020 Feb 1;60:101602. https://doi.org/10.1016/j.media.2019.101602
6. Yin S, Zhang Z, Li H, Peng Q, You X, Furth SL, Tasian GE, Fan Y. Fully-automatic segmentation of kidneys in clinical ultrasound images using a boundary distance regression

network. In2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019) 2019 Apr 8 (pp.1741-1744). IEEE. https://doi.org/10.1109/ISBI.2019.8759170
7. Behboodi B, Rivaz H. Ultrasound segmentation using U-Net: learning from simulated data and testing on real data. In2019 41st annual international conference of the IEEE engineering in medicine and biology society (EMBC) 2019 Jul 23 (pp. 6628-6631). IEEE. https://doi.org/10.1109/EMBC.2019.8857218
8. Zhang Y, Ying MT, Yang L, Ahuja AT, Chen DZ. Coarse-to-fine stacked fully convolutional nets for lymph node segmentation in ultrasound images. In2016 IEEE International Conference on Bioinformatics and Biomedicine (BIBM) 2016 Dec 15 (pp. 443-448). IEEE https://doi.org/10.1109/BIBM.2016.7822557
9. Wu L, Xin Y, Li S, Wang T, Heng PA, Ni D. Cascaded fully convolutional networks for automatic prenatal ultrasound image segmentation. In2017 IEEE 14th international symposium on biomedical imaging (ISBI 2017) 2017 Apr 18 (pp. 663-666). IEEE. https://doi.org/10.1109/ISBI.2017.7950607
10. Kim S, Jang Y, Jeon B, Hong Y, Shim H, Chang H. Fully automatic segmentation of coronary arteries based on deep neural network in intravascular ultrasound images. InIntravascular Imaging and Computer Assisted Stenting and Large-Scale Annotation of Biomedical Data and Expert Label Synthesis: 7th Joint International Workshop, CVII-STENT 2018 and Third International Workshop, LABELS 2018, Held in Conjunction with MICCAI 2018, Granada, Spain, September 16, 2018, Proceedings 3 2018 (pp. 161-168). Springer International Publishing. https://doi.org/10.1007/978-3-030-01364-6_18
11. Mishra D, Chaudhury S, Sarkar M, Soin AS. Ultrasound image segmentation: a deeply supervised network with attention to boundaries.IEEE Transactions on Biomedical Engineering.2018 Oct 21;66(6):1637-48 https://doi.org/10.1109/TBME.2018.2877577
12. Shareef B, Xian M, Vakanski A. Stan: Small tumor-aware network for breast ultrasound image segmentation. In2020 IEEE 17th international symposium on biomedical imaging (ISBI) 2020 Apr 3 (pp. 1-5). IEEE. https://doi.org/10.1109/ISBI45749.2020.9098691
13. Chen G, Dai Y, Li R, Zhao Y, Cui L, Yin X. SDFNet: Automatic segmentation of kidney ultrasound images using multi-scale low-level structural feature. Expert Systems with Applications. 2021 Dec 15;185:115619. https://doi.org/10.1016/j.eswa.2021.115619
14. Chen LC, Zhu Y, Papandreou G, Schroff F, Adam H. Encoder-decoder with atrous separable convolution for semantic image segmentation. In Proceedings of the European conference on computer vision (ECCV) 2018 (pp. 801-818). https://doi.org/10.1007/978-3-030-01234-2_49
15. He K, Zhang X, Ren S, Sun J. 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
16. Russakovsky O, Deng J, Su H, Krause J, Satheesh S, Ma S, Huang Z, Karpathy A, Khosla A, Bernstein M, Berg AC. Imagenet large scale visual recognition challenge. International journal of computer vision. https://doi.org/10.1007/s11263-015-0816-y
17. Singla R, Ringstrom C, Hu G, Lessoway V, Reid J, Nguan C, Rohling R. The open kidney ultrasound data set. InInternational Workshop on Advances in Simplifying Medical Ultrasound 2023 Oct 1(pp. 155-164). Cham: Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-44521-7_15
18. Zhang Q, Cui Z, Niu X, Geng S, Qiao Y. Image segmentation with pyramid dilated convolution based on ResNet and U-Net. InNeural Information Processing: 24th International Conference, ICONIP 2017, Guangzhou, China, November 14-18, 2017, Proceedings, Part II 24 2017 (pp. 364-372). Springer International Publishing. https://doi.org/10.1007/978-3-319-70096-0_38
19. Xia KJ, Yin HS, Zhang YD. Deep semantic segmentation of kidney and space-occupying

lesionarea based on SCNN and ResNet models combined with SIFT-flow algorithm. Journal of medical systems. 2019 Jan;43(1):2. https://doi.org /10.1007/s10916-018-1116-1
20. Xu W, Fu YL, Zhu D. ResNet and its application to medical image processing: Research progress and challenges. Computer Methods and Programs in Biomedicine. 2023 Jun 8:107660. https://doi.org/10.1016/j.cmpb.2023.107660
21. Hasan D, Abdulazeez AM. Lung Segmentation from Chest X-Ray Images Using Deeplabv3plus-Based CNN Model. Indonesian Journal of Computer Science. 2024 Feb 20;13(1). https://doi.org/10.33022/ijcs.v13i1.3700
22. Sun M, Lu L, Hameed IA, Kulseng CP, Gjesdal KI. Detecting small anatomical structures in 3D knee MRI segmentation by fully convolutional networks. Applied Sciences. 2021 Dec 28;12(1):283. https://doi.org/10.3390/app12010283
23. Mejri M, Mejri A, Mejri O, Fekih C. Semantic Segmentation and Object Detection Towards Instance Segmentation: Breast Tumor Identification. arXiv preprint arXiv:2108.03287. 2021 Aug 6. https://doi.org/10.48550/arXiv.2108.03287
24. Xue C, Zhu L, Fu H, Hu X, Li X, Zhang H, Heng PA. Global guidance network for breast lesion segmentation in ultrasound images. Medical image analysis. 2021 May 1;70:101989. https://doi.org/10.1016/j.media.2021.101989