



OPEN ACCESS

EDITED BY

Nicola Strisciuglio,
University of Twente, Netherlands

REVIEWED BY

Amir Faisal,
Sumatra Institute of Technology, Indonesia
Virginia Riego Del Castillo,
University of León, Spain
Lidia Talavera-Martinez,
University of the Balearic Islands, Spain

*CORRESPONDENCE

Serestina Viriri
✉ viriris@ukzn.ac.za

RECEIVED 06 June 2023

ACCEPTED 07 August 2023

PUBLISHED 07 September 2023

CITATION

Abdelrahman A and Viriri S (2023) EfficientNet family U-Net models for deep learning semantic segmentation of kidney tumors on CT images. *Front. Comput. Sci.* 5:1235622. doi: 10.3389/fcomp.2023.1235622

COPYRIGHT

© 2023 Abdelrahman and Viriri. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](#). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

EfficientNet family U-Net models for deep learning semantic segmentation of kidney tumors on CT images

Abubaker Abdelrahman and Serestina Viriri *

School of Mathematics, Statistics and Computer Science, University of KwaZulu-Natal, Durban, South Africa

Introduction: Kidney tumors are common cancer in advanced age, and providing early detection is crucial. Medical imaging and deep learning methods are increasingly attractive for identifying and segmenting kidney tumors. Convolutional neural networks have successfully classified and segmented images, enabling clinicians to recognize and segment tumors effectively. CT scans of kidneys aid in tumor assessment and morphology study, using semantic segmentation techniques for pixel-level identification of kidney and surrounding anatomy. Accurate diagnostic procedures are crucial for early detection of kidney cancer.

Methods: This paper proposes an EfficientNet model for complex segmentation by linking the encoder stage EfficientNet with U-Net. This model represents a more successful system with improved encoder and decoder features. The Intersection over Union (IoU) metric quantifies model performance.

Results and Discussion: The EfficientNet models showed high IoU_Scores for background, kidney, and tumor segmentation, with mean IoU_Scores ranging from 0.976 for B0 to 0.980 for B4. B7 received the highest IoU_Score for segmenting kidneys, while B4 received the highest for segmenting tumors. The study utilizes the KiTS19 dataset for contrast-enhanced CT images. Using Semantic segmentation for EfficientNet Family U-Net Models, our method proved even more reliable and will aid doctors in accurate tumor detection and image classification for early diagnosis.

KEYWORDS

kidney tumor, EfficientNet, deep learning, U-Net, semantic segmentation

1. Introduction

Many factors contribute to the progressive rise in cancer incidence in humans, where effective disease prevention requires early identification and treatment. One of the essential organs in human physiology is the kidney. Kidney cancer requires a precise diagnosis and structured care (Rajinikanth et al., 2023). One of the top 10 cancers that affect both men and women is kidney cancer. The lifetime risk of developing kidney cancer is around 1 in 75 (1.34%). Kidney cancer ranks ninth in men and fourteenth in women. If kidney cancer is detected and treated in its early stages, afflicted individuals have a far higher chance of being cured (Hsiao et al., 2022). The human body has two kidneys and bean-shaped organs on either side of the spine. The kidney plays a crucial role in maintaining body fluid and solute balance through the excretion and filtration of waste products.

It also secretes a variety of hormones and aids in blood pressure regulation. A kidney tumor develops when kidney cells stop functioning and begin to grow quickly. Renal cell carcinoma, which develops from kidney cells, can spread slowly or aggressively. Renal cell cancer frequently manifests as a single tumor. However, any kidney may develop a variety of malignancies (Geethanjali and Dinesh, 2021; Abdelrahman and Viriri, 2022).

Kidney cancer diagnosis usually does not begin with palpation. In most cases, kidney cancer is asymptomatic in its early stages and is often detected incidentally on imaging studies performed for unrelated reasons. Palpation of the kidneys during a physical exam may detect an enlarged kidney, but it is unreliable for detecting kidney cancer (Abdelrahman and Viriri, 2022; Bapir et al., 2022). Once kidney cancer is suspected based on imaging studies, a diagnosis is typically confirmed with a biopsy or surgical tumor resection. A biopsy involves taking a small sample of the tumor tissue for examination under a microscope, while surgical resection involves removing the entire tumor for examination; additional tests may also be performed to determine the extent of cancer and guide treatment decisions. These may include blood tests to assess kidney function and identify any markers of cancer, such as elevated levels of certain proteins, and imaging studies, such as CT or MRI scans, to evaluate the size and location of the tumor and identify any metastases. Follow-up tests include urine analysis, general examinations, lymph node palpations every 2 years, and cystoscopies every 10 years to check for the absence of renal cell carcinoma recurrence (Beisland et al., 2006; Üreyen et al., 2015; Abdelrahman and Viriri, 2022; Bapir et al., 2022).

Manual segmentation methods have been used extensively to study renal carcinoma. Manual kidney segmentation manually outlines the kidney region in medical images by outlining the kidney region and identifying the location and extent of tumors. This is an important step in many clinical applications. Manual kidney segmentation can be time-consuming and subject to inter-observer variability, meaning that different radiologists may segment the kidneys differently. However, it is still widely used as a gold standard for evaluating the accuracy of automated segmentation algorithms. By comparing the results of a machine learning algorithm with manual segmentation, researchers can determine the algorithm's performance and identify areas for improvement (Torres et al., 2018; Abdelrahman and Viriri, 2022). As a result, renal segmentation is an important stage in computer-aided analysis and is particularly important for early cancer detection. In addition to carcinoma, the exact segmentation provides structural evidence of variations in the size and form of the kidneys and tumors, which professionals may utilize to examine significant disorders. Because of the importance of this field to medicine, several researchers choose to employ image processing, machine learning, and deep learning methodologies to widen computational tools to aid experts in the interpretation of clinical images (Hatipoglu and Bilgin, 2017; Nazari et al., 2021; Abdelrahman and Viriri, 2022).

Deep learning techniques have recently shown promising results. It took a lot of time and effort to build earlier attempts to categorize kidney cancer histology images using manual feature extraction techniques and conventional machine learning algorithms. On the other hand, deep learning techniques automate

this procedure (Gurcan et al., 2009). Hence, early identification is essential to guarantee early diagnosis in kidney cancer candidates, boost treatment effectiveness, and decrease death rates. The medical industry has profited from the development and progress of machine learning, yet there is still much opportunity for improvement.

1.1. Research problem

Given the shortage of available medical experts (Kandel and Castelli, 2020; Munien and Viriri, 2021), the potential for inter-observer variability, and the laborious process of making a diagnosis of kidney cancer, all support the need for an automated system to classify kidney cancer histopathology images accurately. By using advanced computer algorithms and machine learning techniques, such a system could potentially improve the speed and accuracy of cancer diagnosis, leading to earlier detection and better patient outcomes. However, it is important to note that any AI-based medical system must be thoroughly validated and tested to ensure its safety and effectiveness before use in clinical practice. Previous approaches have shown promise in addressing this issue. The research explores eight lightweight EfficientNet family architectures, focusing on optimizing resources while maintaining high accuracies, making them valuable tools for resource optimization tasks. They may achieve comparable results to state-of-the-art approaches while consuming less space and training time. This research aims to answer the capability of EfficientNets to achieve similar results to state-of-the-art approaches for classifying kidney cancer CT images (Aresta et al., 2019; Tan and Le, 2019; Abdelrahman and Viriri, 2022). In the next part, we will explore a selection of scholarly papers authored by teachers, focusing specifically on the field of kidney and kidney tumor segmentation.

1.2. Literature review

Computer-aided diagnosis (CAD) systems assist healthcare professionals in making diagnoses by providing additional information and analysis based on medical imaging and data. They use machine learning, deep learning, and image processing to analyze and interpret patient data, emphasizing areas of concern and diagnoses. However, CAD systems are not meant to replace healthcare experts but rather to supplement their knowledge and aid decision-making. Medical practitioners must ultimately diagnose patients and choose appropriate therapy courses based on their clinical judgment and expertise (Ramadan, 2020; Munien and Viriri, 2021; Abdelrahman and Viriri, 2022; Tsuneki, 2022).

1.2.1. Modern approaches

Recent growth in processing power has led to advancements in deep learning-based technologies, particularly CNNs, in medical image processing. These technologies have been successful in segmentation detection, abnormality classification, and retrieval, resulting in the emergence of intriguing algorithms in this field

(Xie et al., 2020). In Ali et al. (2007), the authors use contour templates to describe the shapes of local objects, and Contour templates describe the shapes of local objects like organs, in medical images. These pre-defined shapes capture specific properties of different organs, making them flexible and efficient for segmenting organs without complex models or training data. They can be customized to capture specific features, making them ideal for accurate segmentation in medical imaging for diagnosis and treatment planning (Thong et al., 2018). 2D CNNs are used in medical imaging to segment CT images using two common deep learning segmentation strategies: end-to-end segmentation and cascade segmentation. End-to-end segmentation uses a single deep-learning model, making it simpler and more efficient. However, it may have reduced flexibility, operability, and interoperability, and may require more training data. Cascade segmentation divides the task into multiple stages, using a separate deep-learning model for each stage. Both approaches are widely used in medical image segmentation, depending on the task's requirements, image data complexity, and available computational resources (Xie et al., 2020). Deep Learning focuses on artificial neural networks with multiple layers, used in natural language processing, gaming, image and audio recognition, and assessing visual images (Valueva et al., 2020). CNNs replace matrix multiplication in image processing and recognition, focusing on pixel data (Kim, 2019). They are utilized in image recognition, recommendation systems, classification, segmentation, medical analysis, and more (Collobert and Weston, 2008; Tsantekidis et al., 2017; Avilov et al., 2020).

1.2.2. Convolutional neural network approach

Convolutional neural networks (CNNs) are highly effective learning algorithms for image segmentation, classification, detection, and retrieval tasks (Ciregan et al., 2012; Mayer et al., 2017). CNNs are efficient for image recognition due to their hierarchical representations and spatial structure, making them ideal for categorizing histopathology photos in general computer vision applications (Khan et al., 2020). CNNs have been employed in several studies to identify and categories kidney cancers. The segmentation of kidneys and tumors in abdominal 3D CT scans is the specific emphasis of this section's discussion of several deep-learning techniques for the semantic segmentation of medical images. In the image (G-A), the yellow color indicates the kidney. In the image (G-B), the yellow color indicates Tumor and the green color indicates Kidney. Figure 1 depicts the segmentation of kidneys and tumors, highlighting their distinct regions. Images were taken from KiTS 19 dataset.

1.2.2.1. One-stage methods

One-stage approaches are object identification models that accurately forecast bounding boxes and class probabilities for each item in an image in a single forward pass. Convolutional neural networks (CNNs) use one-stage techniques to extract information from input images, produce feature maps, and predict object bounding boxes and class probabilities. Researchers use deep learning architectures like U-Net, LinkNet-34, and RAU-Net, and improve performance by combining data from multiple scales or training segmentation networks with non-squared patches.

Moreover, some techniques use transformers to record distant relationships for precise tumor segmentation. Some research examples and findings are shown below: The authors (Isensee and Maier-Hein, 2019) propose a "Reset U-Net"-based U-Net architecture for 3D medical image segmentation, addressing the vanishing gradient problem and improving performance. They suggest incorporating a "reset block" to facilitate gradient movement across the network. The authors Myronenko and Hatamizadeh (2019) suggest an automated strategy for segmenting kidney tumors in contrast-enhanced CT images based on deep learning.

Authors use a fully convolutional neural network to semantically segment kidney and tumor areas using KiTS19 challenge dataset contrast-enhanced CT images. The authors Geethanjali and Dinesh (2021) a novel Attention U-Net model used an attention mechanism and modified U-Net architecture to segment kidney tumors from CT data with 0.86 accuracies (Efremova et al., 2019). The research proposes an automated segmentation method for locating kidney and liver cancers in CT scans using a fully convolutional neural network architecture. The method accurately segments kidney and liver tumors, with an average Dice coefficient of 0.76 for kidney and 0.80 for liver tumors (Yang et al., 2018). The authors utilized a weighted loss function in network training to improve segmentation in CT images, potentially aiding radiologists in identifying and treating kidney and renal cancers. Methods include strategic sampling, weighted loss function, and transformers. The Table 1 displays the outcomes of the one-stage methods. All papers in one-stages use the KiTS19 dataset with CT images.

1.2.2.2. Two-stage methods

The two-stage approach in medical image segmentation addresses foreground/background imbalance, improving accuracy and efficiency by detecting the volume of interest (VOI) and segmenting target organs. Variations have strengths and weaknesses depending on application requirements.

Studies on two-stage kidney and tumor segmentation techniques in CT images have been conducted in several instances. da Cruz et al. (2020) created a technique that uses deep (CNN), and it managed to distinguish kidneys with up to 93.03% accuracy. Zhang et al. (2019) developed a cascaded two-stage kidney and tumor segmentation framework that uses a 3D fully convolutional network. Hou et al. (2020) provided a three-stage self-guided network for segmenting kidney tumors that finds the VOI using down-sampled CT images and extracts the kidney and tumor boundaries inside the VOI using the full-resolution net and tumor refine net from full-resolution CT images (Hatamizadeh et al., 2020). Module, which can be used with any general encoder-decoder architecture, improved the edge representations in learnt feature maps (Zhao et al., 2020) created the MSS U-Net, a multi-scale supervised 3D U-Net-based model for segmenting kidney and kidney tumors from CT images (Xie et al. (2020) presented a cascaded SE-ResNeXT U-Net, and (Chen and Liu, 2021) provided a method based on a multi-stage stepwise refinement strategy for segmenting kidney, tumor, and cyst in the abdomen-enhanced CT images. Finally, Wen et al. (2021) proposed the SeResUNet segmentation network, which leverages ResNet to deepen the encoder's network and speed convergence and is specifically

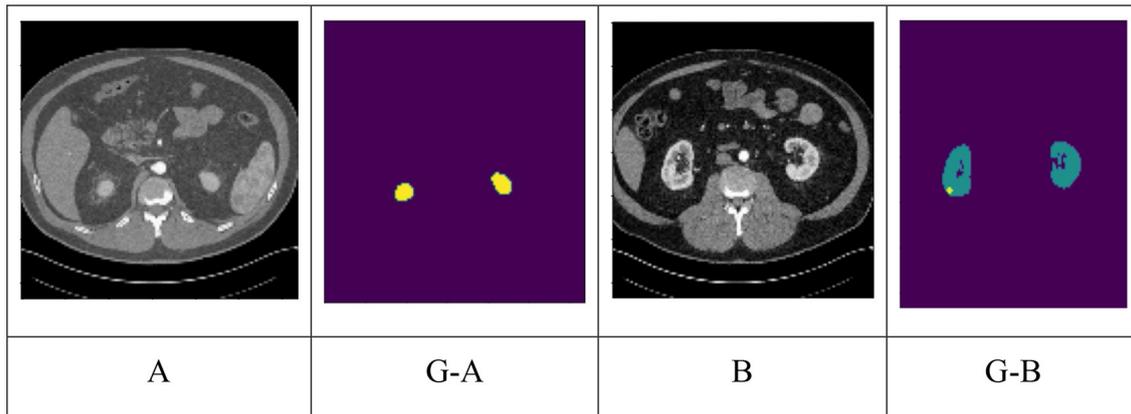


FIGURE 1 Different classes of kidney dataset (A) normal, (G-A) ground truth of (A). Yellow indicates kidney; (B) tumor; (G-B) ground truth of (B). Green indicates kidney and Yellow indicates tumor.

TABLE 1 Displays the outcomes of several one-stage methods.

Reference	Input	Regulation	Architecture	Loss	Optimizer	Kidneys dice	Tumor dice	Kidneys IoU scores	Tumor IoU scores
Isensee and Maier-Hein (2019)	3D		3D U-Net	Dice	SDG	0.973	0.850	-	-
Santini et al. (2019)	2.5D	BN	mult-istage-U-Net	Dice	Adam	0.96	0.74	-	-
Causey et al. (2021)	2D		Ensemble	Dice		0.947	0.778	-	-
Myronenko and Hatamizadeh (2019)	3D		Boundary-Aware Architecture	KD, TD, CD	Adam	0.974	0.810	-	-
Guo et al. (2021)	3D		RAU-Net	Dice SDG	Dice SDG	0.960	0.770	-	-
Efremova et al. (2019)	2D,3D		CNN	Dice		0.964	0.674	-	-
Geethanjali and Dinesh (2021)	3D		U-Net	IoU_Scores		-	-	0.957	0.939

designed to segment the kidney and tumor. Regarding kidney and tumor segmentation, two-stage approaches have generally shown encouraging results, but there is always potential for improvement, particularly for smaller organs and tumors. Table 2 displays the outcomes of the Two-Stage Methods. All papers in two-stages use the KiTS19 dataset with CT images except (Wen et al., 2021) use KiTS21 with CT images.

Two-stage medical image segmentation technique detects Voids and targets organs using deep convolutional neural networks, image processing methods, and architectures. Improvements are needed for smaller kidneys, tumors, and cysts. In kidney tumor semantic segmentation, one-stage methods offer a simpler and more efficient approach by employing a single neural network for end-to-end segmentation. This results in faster inference times and makes them well-suited for real-time applications. However, these methods may compromise accuracy and stress to capture fine details and handle complex structures. On the other hand, two-stage methods present a more sophisticated strategy with separate proposal generation and segmentation

refinement steps, leading to higher accuracy, robustness in handling challenging cases, and the ability to support hierarchical processing and feature reuse. Nevertheless, this increased accuracy comes at the cost of greater computational complexity, and they tend to have longer inference times, making them less ideal for real-time scenarios compared to their one-stage counterparts.

1.3. Research contributions

The study evaluated eight EfficientNets versions for kidney cancer CT image classification. The architecture effectively extracted and learned global image features, including tissue and nuclei organization. EfficientNet-B4 and EfficientNet-B7 models showed the best performance, with an accuracy mean IoU of 0.980. EfficientNet-B4 was superior for tumor detection and EfficientNet-B7 for kidney detection. The study highlights the potential of EfficientNets for kidney cancer and CT image classification due to their simplicity, reduced training time, and consistent accuracies.

TABLE 2 The outcomes of several two-stage methods.

Reference	Input	Regulation	Architecture	Loss	Optimizer	Kidneys dice	Tumor dice	Kidneys IoU scores	Tumor IoU scores
Zhang et al. (2019)	3D	BN	Cascaded	Dice	Adam	0.974	0.831	-	-
Hou et al. (2020)	3D		U-Net	Dice	Adam	0.967	0.845	-	-
Hatamizadeh et al. (2020)	3D		EG-CNN	Dice	Adam	0.965	0.835	-	-
Xie et al. (2020)	3D	BN	Cascaded	Dice	Adam	0.967	0.743	-	-
Wen et al. (2021)	3D		SeResUNet	Dice	Adam	0.916	0.542	-	-

1.4. Paper structure

This paper is organized into different sections. Part 2, Material and Method, the methods and techniques section, provides information on the framework used in this study. Part 3, Results and Discussion, is dedicated to presenting the results obtained during the research. Finally, Part 4, Conclusion, discusses the insights gained from this study and concludes the paper.

2. Material and method

The dataset used in this study is KiTS19, which stands for the Kidney Tumor Segmentation Challenge 2019. KiTS19 is a publicly available dataset of CT scans of the abdomen with kidney tumors (Heller et al., 2019). Evaluation Metrics: The evaluation metric used in this study is the Intersection over Union (IoU), also known as the Jaccard index. IoU is a common evaluation metric used for object detection and segmentation tasks.

2.1. Dataset

The test facts are gathered from KiTS19 (Heller et al., 2019). It is a dataset for kidney tumor segmentation that includes CT scans and corresponding annotations for training and testing machine learning models. It is a widely used benchmark dataset for evaluating the performance of medical image segmentation models on kidney tumors. Two hundred ten abdominal CT volumes comprise the testing and challenging training datasets. All slices are seen axially, and each volume's imaging and ground truth labels are supplied in Neuroimaging Informative Technology Informative (DICOM) format. This can be found on Kaggle at <https://www.kaggle.com/competitions/body-morphometry-kidney-and-tumor/data> (Who used the original data in Heller et al., 2019) with shape (number of slices, height, and breadth). The dataset comprises slices between 29 and 1059 pixels and grayscale images with a size of 512 × 512 pixels. The following Table 3, lists the characteristics of the experimental dataset. We use part of the data, 7,899 png images.

TABLE 3 Dataset properties used for the experimentation.

Properties	Values
Number of all image	7,899
Number of training images	5,841
Number of validation images	1,027
Number of testing images	1,031
Image format	Png
Modality	CT

2.2. Preprocessing

Preprocessing is indeed crucial for the classification of images. Images are typically stained with various dyes to highlight different structures and features, and they can vary significantly in terms of staining intensity, color, and texture. Therefore, preprocessing is necessary to enhance the images' quality and consistency and remove noise and artifacts that may interfere with the classification process (Aresta et al., 2019). They are rather large, while convolutional neural networks are typically designed to take in much smaller inputs. Therefore, the resolution of the images must be decreased so that the network can receive the input while maintaining the important features. In addition to the preprocessing steps mentioned earlier, another important step is to resize or down sample the CT images to a suitable size for processing by convolutional neural networks (CNNs). CNNs are typically designed to take in images of a fixed size, and the size of the input images can significantly impact the network's performance. Therefore, the resolution of the CT images must be decreased while preserving the important features and structures in the image. This can be achieved using down-sampling techniques, such as bilinear or nearest-neighbor interpolation, or more advanced techniques, such as wavelet-based or deep learning-based methods.

2.3. U-Net

U-Net is a convolutional neural network (CNN) architecture used for image segmentation tasks, which was proposed by Olaf Ronneberger, Philipp Fischer, and Thomas Brox in 2015. The network is named after its U-shape architecture, which consists

of a contracting path (encoder) and an expansive path (decoder). The contracting path consists of convolutional and max pooling layers, which reduce the spatial resolution of the input image while increasing the number of feature channels. This process enables the network to capture the high-level features of the image. The expansive path consists of transposed convolutional layers, which increase the spatial resolution of the feature maps while decreasing the number of feature channels. This path enables the network to perform pixel-wise classification by predicting a binary mask for each pixel in the input image. The U-Net architecture has become popular in the medical imaging community for its ability to segment organs and other structures from medical images accurately. It has also been used for other applications, such as the segmentation of objects in satellite images and the semantic segmentation of urban scenes (Ronneberger et al., 2015). The U-Net is a convolutional neural network for biomedical image segmentation. Its architecture, consisting of an encoder-decoder design, is highly stable and can achieve precise segmentation with fewer training images. The network comprises 3×3 convolutional layers, with a maximum of 2×2 following each pooling layer and the ReLU activation function. A 1×1 convolutional layer is attached at the end. The U-Net consists of path contracting and symmetric expanding, used for capturing and precise localization. The U-Net architecture depends heavily on data augmentation techniques. The network can segment a 512×512 image on a modern GPU in less than a second. The U-Net has been successfully applied to medical image segmentation, although 3D convolution is suggested for fully utilizing the spatial information of 3D images such as CT and MRI. Based on the 3D U-Net, Fabian made minor changes and achieved first place in several medical image segmentation contests, demonstrating that an optimized U-Net can outperform many innovative designs (Ronneberger et al., 2015; Payer et al., 2016).

2.4. EfficientNet

Convolutional neural network (CNN) researchers at Google AI unveiled the EfficientNet architecture in Figure 2. The primary objective of EfficientNet was to develop a high-accuracy and computationally efficient CNN that could be trained and used on various platforms, including mobile phones and other systems with limited resources. Compound scaling, which involves scaling each of the neural network's dimensions (such as depth, breadth, and resolution) in a balanced manner, is the foundation of the design of EfficientNet. This enables higher performance and efficiency than conventional scaling techniques, which often concentrate on just one or two network dimensions. EfficientNet uses a cutting-edge "compound coefficient optimization" technique to balance depth, breadth, and resolution scaling. This technique is one of the main inventions of EfficientNet. This enables higher accuracy levels while making better use of computer resources. On a variety of computer vision tasks, such as image classification, object recognition, and semantic segmentation, EfficientNet has produced state-of-the-art results. For instance, EfficientNet outperformed all other models in the ImageNet Large Scale Visual Recognition Challenge (ILSVRC) 2019 with a top-1 accuracy of 88.4% and a top-5 accuracy of 97.4%.

EfficientNet has been demonstrated to be extremely resource-efficient and have great accuracy. For instance, EfficientNet reduced parameter count by ten times and inference time by 2.8 times in one research compared to the prior state-of-the-art model on a mobile device. In both business and academics, EfficientNet is widely utilized for various purposes, including augmented reality, medical imaging, and autonomous driving. Researchers and developers may experiment with and modify the architecture because it is open-source (Tan and Le, 2019, 2021).

EfficientNet is a powerful convolutional neural network (CNN) architecture that has shown impressive results on a wide range of computer vision tasks. One such task is medical image analysis, where EfficientNet has been used to analyze medical images for various purposes, including cancer detection.

2.5. Evaluation metrics

The study used two evaluation metrics for evaluating image segmentation performance: IoU (Intersection over Union) and Dice similarity coefficient (DSC). IoU quantifies the similarity between predicted and ground truth masks by calculating the ratio of common pixels. DSC, also known as F1 score or Sørensen-Dice index, is a balanced measure of segmentation accuracy. Both metrics are appropriate for evaluating segmentation results, but the IoU is more sensitive to under- and over-segmentation errors. Despite this, DSC is the most commonly used metric in scientific publications for medical image segmentation evaluations. These metrics are calculated using the formulae shown below.

$$\text{IoU} = \frac{\text{True Positive}}{(\text{True Positive} + \text{False Positive} + \text{False Negative})} \quad (\text{Pandey et al., 2020}) \quad (1)$$

$$\text{DSC} = \frac{2\text{TP}}{2\text{TP} + \text{FP} + \text{FN}} \quad (\text{Pandey et al., 2020}) \quad (2)$$

2.6. Methods and techniques

The methodology presented in the study aims to segment tumors and kidneys on CT slices using the KiTS19 database (Heller et al., 2019). The method consists of two main stages: data preparation and segmentation using an EfficientNet U-Net model. In the data preparation stage, the CT volumes are first scaled and normalized to improve the consistency and quality of the data. Then, the volumes are transferred from DICOM to png format for further processing. The segmentation stage involves using an EfficientNet U-Net model to segment the kidneys and the tumor concurrently. The U-Net model is a type of convolutional neural network that is widely used for image segmentation tasks. The EfficientNet variant of the U-Net model is a more efficient and accurate version of the standard U-Net, which uses fewer parameters and achieves better results. Preprocessing is also essential for the methodology presented in the study involves data preparation and segmentation using an EfficientNet U-Net model.

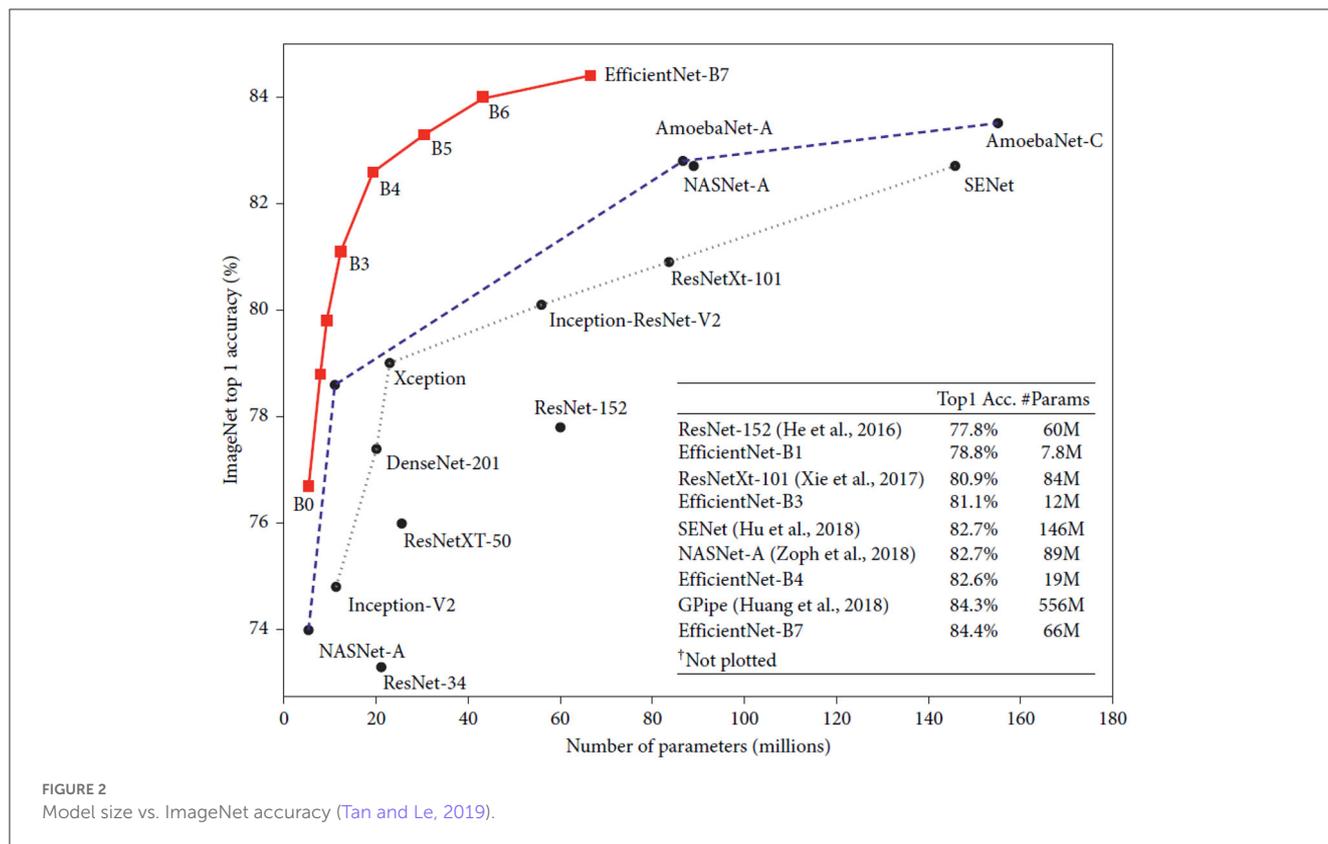


FIGURE 2 Model size vs. ImageNet accuracy (Tan and Le, 2019).

2.7. Methodology

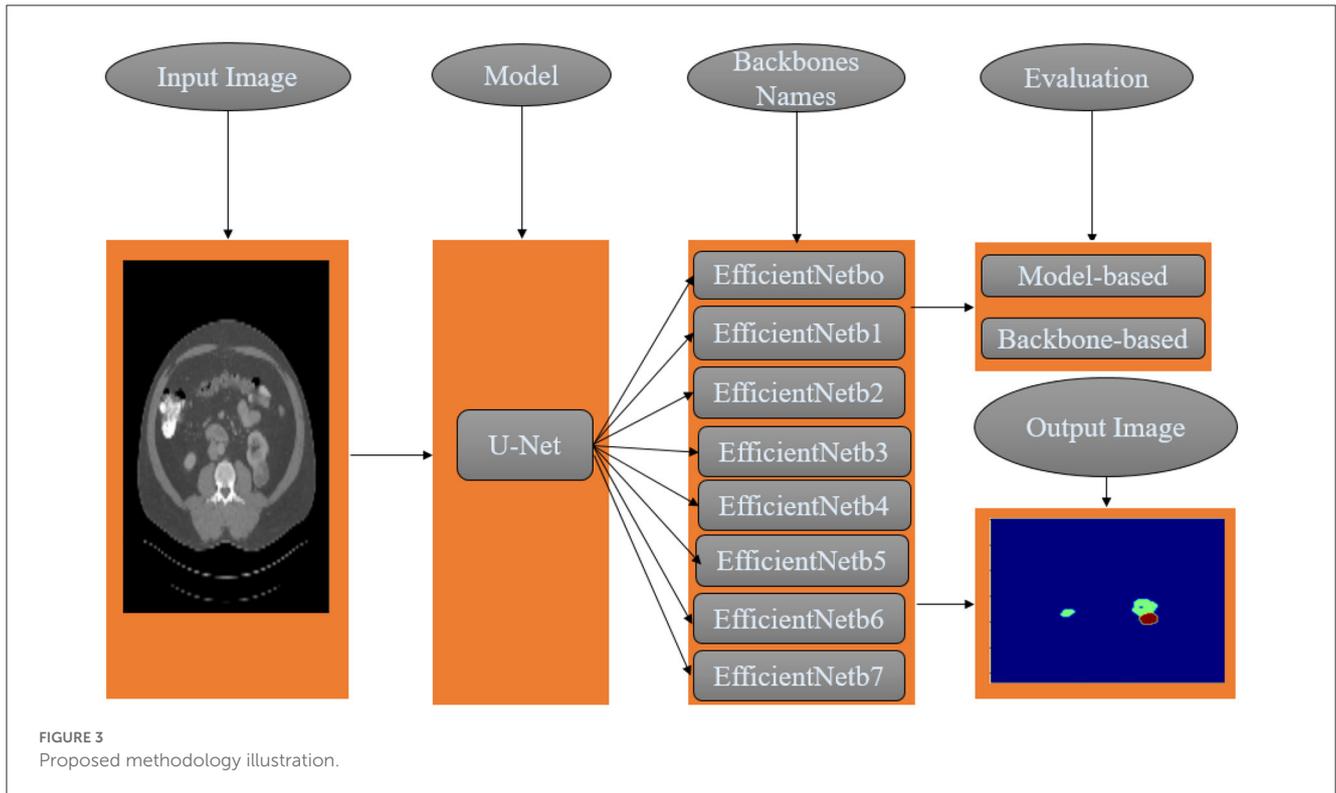
This section explains the suggested method for segmenting tumors and kidneys on CT slices. Slices of CT were obtained from the KiTS19 database to conduct the studies (Heller et al., 2019). Each volume received data preparation in the first stage, which comprised scaling, normalization, and transferring CT volumes from DICOM to png format. Using the EfficientNet U-Net model in the segmentation stage, the kidneys and tumor are segmented concurrently in the second step. The second step of the pipeline involves using an EfficientNet U-Net model to simultaneously segment the kidneys and tumors. The U-Net architecture is commonly used in medical image segmentation tasks because it effectively captures local and global image features. The EfficientNet variant of the U-Net model is designed to be computationally efficient, which makes it suitable for processing large volumes of medical imaging data. Finally, the segmentation results for the kidneys and tumors are obtained from the ensemble phase, where multiple predictions from the EfficientNet U-Net model are combined to produce the best possible segmentation. This is typically done by computing the intersection-over-union (IoU) scores for the different predictions and selecting the segmentation with the highest IoU score.

The IoU_Score measures the overlap between the predicted and ground truth segmentation and is commonly used to evaluate segmentation performance in medical imaging. Segmentation models: this is an open-source library of deep-learning models

for image segmentation tasks. The library implements state-of-the-art architectures for semantic, instance, and panoptic segmentation. Some of the models available in the library include UNet, PSPNet, FPN, LinkNet, and MaskRCNN, among others. The library is built on the Keras deep learning framework and supports TensorFlow and PyTorch backends. It also includes pre-trained weights for various datasets such as Cityscapes, Pascal VOC, and COCO, which can be used for transfer learning or fine-tuning on custom datasets. The segmentation models library is useful for researchers and practitioners working on image segmentation tasks. This can be found on GitHub at https://github.com/qubvel/segmentation_models.

2.8. Backbone architectures

The backbones of feature extraction networks compute image input features, and selecting the optimal network is crucial for objective task performance and Deep Learning (DL) model computational complexity. Numerous backbone networks have been designed and implemented in various DL models. Further research is needed to compare feature extraction networks for DL applications (Elharrouss et al., 2022). This study evaluates eight existing feature extraction backbone networks for a single U-Net model to determine the most effective combination. Unsuitable backbones can degrade performance, be computationally expensive, and be complex.



2.9. Proposed methodology

This study applies semantic segmentation to renal tumors. The proposed methodology combines one semantic segmentation U-Net model and eight feature extraction networks. This modular architecture seeks to identify the optimal segmentation solution for kidney tumors. In addition, the optimal Loss function is examined and evaluated from both the model and backbone perspectives, and a methodology is proposed as shown in Figure 3.

2.10. Implementation

All algorithms in this study were implemented in Python 3.9.16 using Anaconda and Jupyter on a DELL CORE i5 personal computer. Original images (512 × 512) were resized uniformly to 128 × 128 pixels for segmentation model input. This section discusses assembling the eight models EfficientNet family “EfficientNetB0”, “EfficientNetB1”, “EfficientNetB2”, “EfficientNetB3”, “EfficientNetB4”, “EfficientNetB5”, “EfficientNetB6”, and “EfficientNetB7”, using U-Net architecture.

2.11. Model setup

Each model is trained for a total of 50 epochs. To avoid overfitting, training ceases when the validation loss remains at or above 0.0001. All backbones have ImageNet-trained weights for accelerated convergence. Table 4 details all DL models’ baseline model configurations and hyperparameters.

TABLE 4 Models’ hyperparameter setup.

Hyperparameter	Settings
Activation	“softmax”
Optimizer	Adam
Learning rate	0.0001
Batch size	32
Epochs	50
Metrics	IoU
Input images size	128 × 128

3. Results and discussion

This section discusses assembling the eight models EfficientNetB0, EfficientNetB1, EfficientNetB2, EfficientNetB3, EfficientNetB4, EfficientNetB5, EfficientNetB6, and EfficientNetB7 using U-Net architecture. To train all models using EfficientNet U-Net architecture.

3.1. Experimental results

The accuracy of each model using the stain normalization techniques is shown in Table 5, along with the average precision and recall (over the three classes) for each model. Additionally, the EfficientNet model’s average accuracy is determined. In this activity, accuracy and sensitivity (precision) is crucial. As

TABLE 5 Outcomes of the ensemble model of EfficientNetB7.

IoU_Scores	Obtained values
Background IoU_Score	0.999
Kidney IoU_Score	0.977
Tumor IoU_Score	0.962
Mean IoU_Score	0.980

a result, these indicators will be examined and explored in more detail.

3.2. Discussion

3.2.1. Analysis of results

Outcomes revealed that the best-performing models were pertained on ImageNet and fine-tuned. Even though many approaches incorporate patch-wise classification to utilize local features (Aresta et al., 2019), using the entire image for classification produced better results. This insight implies that extracting nuclei and tissue organization features is more valuable for deciphering the image classes than nuclei-scale features. The image-level classification in this research observes the ability of the architecture to extract global features in kidney cancer CT images and use it to classify unseen images Table 6 shows the result.

Convolutional neural network (CNN) models belonging to the EfficientNet family are intended to be efficient in terms of model size and computing cost while maintaining excellent accuracy. The models are intended to be scalable in complexity and size and are trained on big image datasets. Several models in the EfficientNet family are identified by the letters B0–B7, with B0 being the smallest and B7 being the biggest and most sophisticated. The outcomes shown in Table 6 demonstrate how effectively the EfficientNet designs handle the particular image segmentation tasks assessed in this study. The models obtained high IoU_Scores for background, kidney, and tumor segmentation, with mean IoU_Scores ranging from 0.974 for B1 to 0.980 for B4. The chart demonstrates that performance on the segmentation job typically improves when the EfficientNet architecture version grows from B0 to B7. The Mean IoU_Score, for instance, rises from 0.976 for B0 to 0.980 for B4, then slightly falls for B5, B6, and B7. The findings demonstrate that the background segmentation IoU_Score was consistently high, with a score of 0.999, across all EfficientNet designs (B0–B7). B7 received the greatest IoU_Score for segmenting kidneys, scoring 0.977, whereas B4 received the highest IoU_Score for segmenting tumors, scoring 0.968. With mean IoU_Scores ranging from 0.976 for B0 to 0.980 for B4, the results indicate that the EfficientNet designs perform well on these segmentation tasks. These results show that the models can accurately distinguish and segment the required areas of interest in the photos. According to the Table 6, the EfficientNet architecture successfully segments images overall, with higher versions often offering greater performance.

The results presented in Table 6 highlight the exceptional performance of the EfficientNet architectures in handling the specific image segmentation tasks evaluated in this study. The

models achieved impressive Dice similarity coefficient (DSC) Scores for background, kidney, and tumor segmentation, with mean Dice Scores ranging from 0.987 for B1 to 0.990 for B4. Notably, the performance generally improved as the EfficientNet architecture version progressed from B0 to B7. For instance, the Dice Score increased from 0.987 for B1 to 0.990 for B4, then slightly declined for B5, B6, and B7. Furthermore, the background segmentation Dice Scores were consistently high, with a score of 0.999, across all EfficientNet designs (B0–B7). Among them, B7 achieved the highest Dice Score of 0.988 for segmenting kidneys, while B4 obtained the top Dice Score of 0.984 for segmenting tumors. These findings indicate that the EfficientNet designs perform exceptionally well in addressing the challenges of these segmentation tasks.

3.2.2. Computational cost data between the different versions EfficientNetB4 and EfficientNetB7

EfficientNetB4 has a smaller model size compared to EfficientNetB7, which has more layers and parameters, making it larger and more complex. The increased model size allows for more complex patterns and features in data. However, EfficientNetB7 takes longer to train due to its increased complexity and number of parameters. The specific computational cost and performance of these models can vary depending on the implementation, hardware setup, and dataset nature. Table 7 shows the number of parameters and time cost of our model.

According to the Table 8, the (Sharma et al., 2020) model achieved an IoU_Score of 0.97 for kidney and 0.32 for tumor segmentation. On the other hand, “Our Model B4” achieved higher scores of 0.974 for kidney segmentation and 0.968 for tumor segmentation. This suggests that “Our Model B4” outperformed the “Reference” model in accurately segmenting the kidney and tumor regions.

3.2.3. Model accuracy and model loss

The EfficientNet model's performance is evaluated in terms of loss and IoU_Score, indicating its ability to minimize predicted and actual values. The model also measures the overlap between predicted and ground truth bounding boxes. Training and validation IoU_Score and training and validation loss are also measured. For example, the model avoids over-fitting and under-fitting to outlier points in Figures 4, 5.

Table 9 shows the performance of different segmentation approaches on the KiTS19 dataset, where each approach is evaluated based on its kidney and tumor scores. The scores range from 0 to 1, with higher scores indicating better segmentation performance. The approaches listed in the table include various types of convolutional neural networks (CNNs), such as U-Net, V-Net, and RAU-Net, as well as cascaded and boundary-aware networks. Additionally, the table includes the scores for a 3D U-Net and a 3D SEAU-Net. The last two rows of the table represent the scores for “Our approach,” which uses U-Net with two different versions of EfficientNet as its backbone. The first version uses EfficientNet-B4, while the second version

TABLE 6 Results for the EfficientNet architectures using mean IoU, and mean dice.

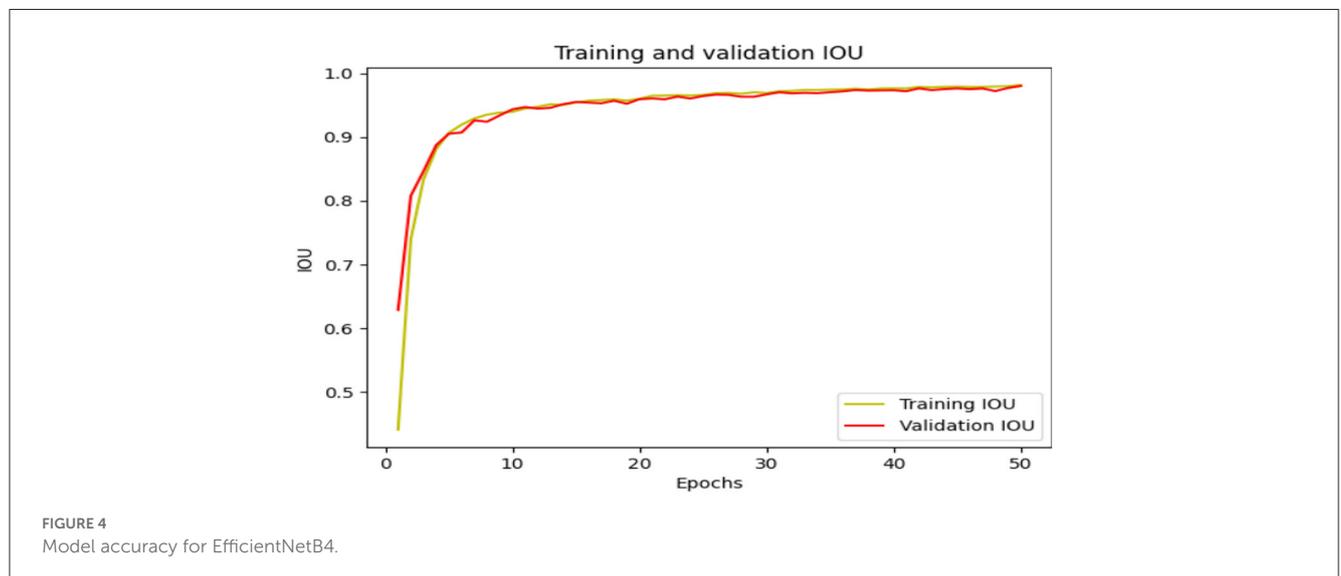
EfficientNet	Background IoU_Score	Kidney IoU_Score	Tumor IoU_Score	Mean IoU_Score	Background dice	Kidney dice	Tumor dice	Mean dice
B0	0.999	0.969	0.960	0.976	0.999	0.984	0.980	0.988
B1	0.999	0.963	0.961	0.974	0.999	0.981	0.980	0.987
B2	0.999	0.966	0.965	0.977	0.999	0.983	0.982	0.988
B3	0.999	0.970	0.960	0.976	0.999	0.985	0.980	0.988
B4	0.999	0.974	0.968	0.980	0.999	0.987	0.984	0.990
B5	0.999	0.972	0.966	0.979	0.999	0.986	0.983	0.989
B6	0.999	0.971	0.963	0.978	0.999	0.985	0.981	0.988
B7	0.999	0.977	0.962	0.980	0.999	0.988	0.981	0.989

TABLE 7 Computational cost data between the different versions EfficientNetB4 and EfficientNetB7.

Versions EfficientNet	Total params	Trainable params	Non-trainable params	Time per second
EfficientNetb4	25.735.307	25.608.123	127.184	114,386
EfficientNetb7	75.048.387	74.735.682	312.704	118,455

TABLE 8 Comparison of object detection results using intersection over union (IoU) with previous studies.

Reference	Dataset	Architecture	Type of images	Kidney IoU_Score	Tumor IoU_Score
Sharma et al. (2020)	KiTS19	U-Net	CT	0.97	0.32
Our Model B4	KiTS19	EfficientNe U-Net	CT	0.974	0.968



uses EfficientNet-B7. The IoU_scores for Our approach are quite high, with a kidney score of 0.974 and 0.977 and a tumor score of 0.968 and 0.962, respectively. These scores suggest that our approach is highly effective for segmenting kidney and tumor regions in the KiTS19 dataset. Figure 6 compares different segmentation approaches.

4. Conclusion

In conclusion, this paper proposes a method for segmenting tumors and kidneys on CT slices using an EfficientNet U-Net model. The suggested process involves data preparation, segmentation, and an ensemble phase. Preprocessing, down

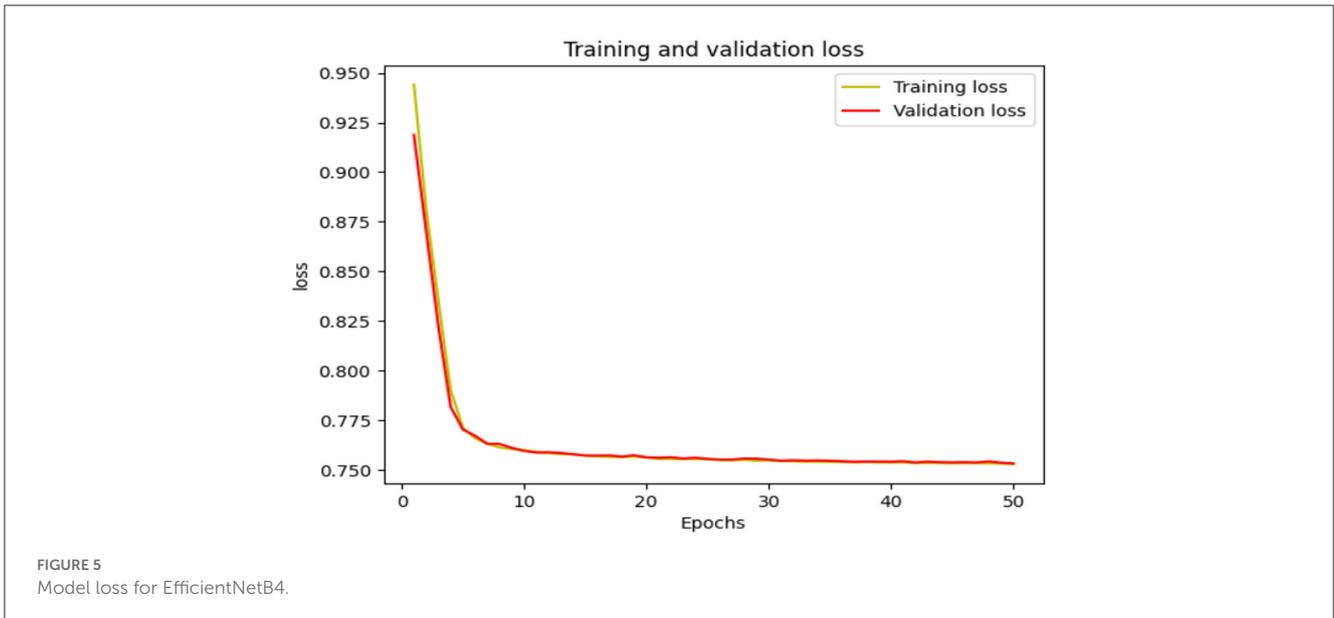
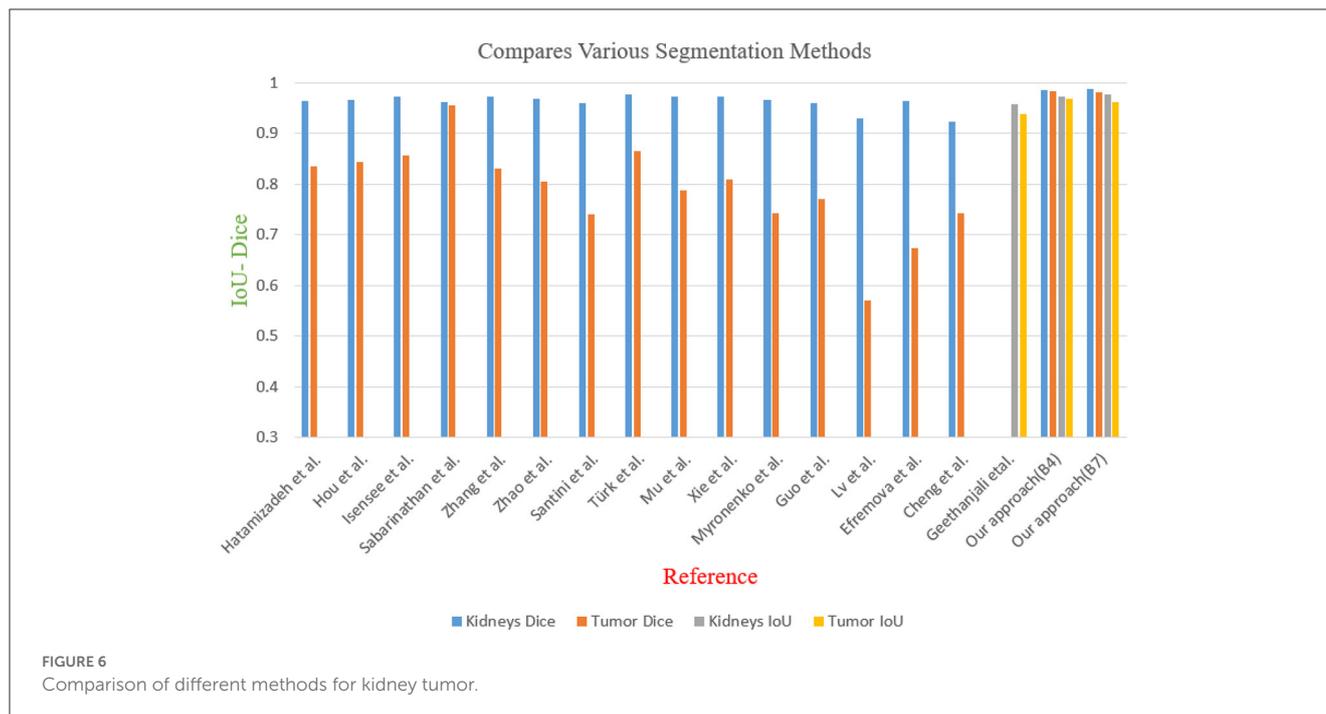


TABLE 9 KiTS19 dataset segmentation: evaluating pre-trained architecture approaches against prior methods.

Reference	Architecture	Kidneys dice	Tumor dice	Kidneys IoU scores	Tumor IoU scores
Hatamizadeh et al. (2020)	CNN	0.965	0.835	-	-
Hou et al. (2020)	U-Net	0.967	0.845	-	-
Isensee and Maier-Hein (2019)	3D U-Net	0.974	0.856	-	-
Sabarinathan et al. (2020)	U-Net	0.963	0.955	-	-
Zhang et al. (2019)	Cascaded	0.974	0.831	-	-
Zhao et al. (2020)	U-Net	0.969	0.805	-	-
Santini et al. (2019)	Multi-stage U-Net	0.96	0.74	-	-
Türk et al. (2020)	V- Net	0.977	0.865	-	-
Mu et al. (2019)	V- Net	0.974	0.789	-	-
Myronenko and Hatamizadeh (2019)	Boundary-Aware	0.974	0.810	-	-
Xie et al. (2020)	Cascaded	0.967	0.743	-	-
Guo et al. (2021)	RAU-Net	0.960	0.770	-	-
Lv and Wang (2021)	3D U-Net	0.930	0.570	-	-
Efremova et al. (2019)	CNN	0.964	0.674	-	-
Cheng et al. (2019)	3D SEAU -Net	0.924	0.743	-	-
Geethanjali and Dinesh (2021)	U-Net	-	-	0.957	0.939
Our approach	U-net EfficientNet-B4	0.987	0.984	0.974	0.968
Our approach	U-net-EfficientNet-B7	0.988	0.981	0.977	0.962

sampling, and resizing CT images are also essential to classifying CT images. The paper also discusses the IoU metric as a common evaluation metric for image segmentation tasks. Finally, the paper summarizes related techniques in the literature and the U-Net architecture used for image segmentation tasks. The proposed method can improve the segmentation accuracy and the diagnosis of tumors and kidneys on CT slices.

This study explored the performance of the U-Net architecture when combined with various models of the EfficientNet family for image segmentation in kidney cancer CT images. The results showed that the models fine-tuned on ImageNet outperformed patch-wise classification approaches regarding accuracy and sensitivity. The EfficientNet models showed excellent background, kidney, and tumor segmentation results, with mean IoU_Scores



ranging from 0.976 to 0.980. In particular, B4 achieved the highest IoU_Score for segmenting tumors, with a score of 0.968, while B7 received the highest IoU_Score for segmenting kidneys, with a score of 0.977. Comparison with previous studies revealed that the models presented in this study achieved higher IoU_Scores for kidney and tumor segmentation. These results suggest that the U-Net architecture combined with EfficientNet models can accurately distinguish and segment the required areas of interest in kidney cancer CT images.

The last table presents the performance of different segmentation approaches for kidney and tumor regions in the KiTS19 dataset, where higher scores indicate better segmentation performance. The table includes several types of convolutional neural networks, such as U-Net, V-Net, and RAU-Net, alongside cascaded and boundary-aware networks. “Our approach,” which utilizes U-Net with two different versions of EfficientNet as its backbone, achieves impressive Dice scores of 0.987 and 0.988 for the kidney and 0.984 and 0.981 for the tumor scores. These results suggest that our approach is a highly effective segmentation technique for the KiTS19 dataset.

Based on the findings and conclusions of this study, there are several potential future directions for research in this area. Some possible areas for future work include further optimization of the proposed method: the proposed method in this paper achieved high accuracy in segmenting tumors and kidneys on CT slices. However, there is still room for further optimization and improvement of the method. For example, the authors could explore using different pre-processing techniques, different data augmentation strategies, or other modifications to the U-Net architecture to improve performance further. Application to other medical imaging datasets: the proposed method was evaluated on the KiTS19 dataset, a benchmark dataset for kidney tumor segmentation. However, it would be interesting to see how well

the method performs on other medical imaging datasets, such as lung, brain, or liver imaging. Comparing the performance of the proposed method with other state-of-the-art techniques on different datasets could help to establish its generalizability and robustness. Exploration of other deep learning architectures: the authors of this paper used the U-Net architecture combined with EfficientNet models for image segmentation tasks. However, many other deep learning architectures, such as Attention U-Net, DeepLab, or Mask R-CNN, could be explored for this purpose. Future work could explore using different architectures and compare their performance with the proposed method.

Data availability statement

Publicly available datasets were analyzed in this study. The original dataset is in <https://kits19.grand-challenge.org/data/>, and we used the dataset found at <https://www.kaggle.com/competitions/body-morphometry-kidney-and-tumor/data>.

Author contributions

All authors listed have made a substantial, direct, and intellectual contribution to the work and approved it for publication.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated

organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

- Abdelrahman, A., and Viriri, S. (2022). Kidney tumor semantic segmentation using deep learning: a survey of state-of-the-art. *J. Imag.* 8, 55. doi: 10.3390/jimaging8030055
- Ali, A. M., Farag, A. A., and El-Baz, A. S. (2007). "Graph cuts framework for kidney segmentation with prior shape constraints," in *Medical Image Computing and Computer-Assisted Intervention—MICCAI 2007: 10th International Conference, Brisbane, Australia, October 29–November 2, 2007, Proceedings, Part I 10* (Berlin Heidelberg: Springer) 384–392. doi: 10.1007/978-3-540-75757-3_47
- Aresta, G., Araújo, T., Kwok, S., Chennamsetty, S. S., Safwan, M., Alex, V., et al. (2019). BACH: Grand challenge on breast cancer histology images. *Med. Image Anal.* 56, 122–139. doi: 10.1016/j.media.2019.05.010
- Avilov, O., Rimbart, S., Popov, A., and Bougrain, L. (2020). "Deep learning techniques to improve intraoperative awareness detection from electroencephalographic signals," in *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS 142–145*. doi: 10.1109/EMBC44109.2020.9176228
- Bapir, R., Hammood, Z. D., Omar, S. S., Salih, A. M., Kakamad, F. H., Najjar, K. A., et al. (2022). Synchronous invasive ductal breast cancer with clear cell renal carcinoma: a rare case report with review of literature. *IJS Short Rep.* 7, e59. doi: 10.1097/SR9.0000000000000059
- Beisland, C., Talleraas, O., Bakke, A., and Norstein, J. (2006). Multiple primary malignancies in patients with renal cell carcinoma: A national population-based cohort study. *BJU Int.* 97, 698–702. doi: 10.1111/j.1464-410X.2006.06004.x
- Causey, J., Stubblefield, J., Qualls, J., Fowler, J., Cai, L., Walker, K., et al. (2021). An ensemble of U-Net models for kidney tumor segmentation with CT images. *IEEE/ACM Trans. Comput. Biol. Bioinform.* 5963, 1–5. doi: 10.1109/TCBB.2021.3085608
- Chen, Z., and Liu, H. (2021). "5 D cascaded semantic segmentation for kidney tumor Cyst," in *International Challenge on Kidney and Kidney Tumor Segmentation* (Cham: Springer International Publishing) 28–34. doi: 10.1007/978-3-030-98385-7_4
- Cheng, J., Liu, J., Liu, L., Pan, Y., and Wong, J. (2019). *A Double Cascaded Framework Based on 3D SEAU-Net for Kidney and Kidney Tumor Segmentation*. University of Minnesota Libraries Publishing. doi: 10.24926/548719.067
- Ciregan, D., Meier, U., and Schmidhuber, J. (2012). "Multi-column deep neural networks for image classification," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition* 3642–3649. doi: 10.1109/CVPR.2012.6248110
- Collobert, R., and Weston, J. (2008). "A unified architecture for natural language processing: Deep neural networks with multitask learning," *Proceedings of the 25th International Conference on Machine Learning* 160–167. doi: 10.1145/1390156.1390177
- da Cruz, L. B., Araújo, J. D. L., Ferreira, J. L., Diniz, J. O. B., Silva, A. C., de Almeida, J. D. S., et al. (2020). Kidney segmentation from computed tomography images using deep neural network. *Comput. Biol. Med.* 123, 103906. doi: 10.1016/j.combiomed.2020.103906
- Efremova, D. B., Kononov, D. A., Siriapisith, T., Kusakunniran, W., and Haddawy, P. (2019). Automatic segmentation of kidney and liver tumors in CT images. *arXiv preprint arXiv:1908.01279*. doi: 10.24926/548719.038
- Elharrouss, O., Akbari, Y., Almaadeed, N., and Al-Maadeed, S. (2022). Backbones-review: Feature extraction networks for deep learning and deep reinforcement learning approaches. *arXiv preprint arXiv:2206.08016*.
- Geethanjali, T. M., and Dinesh, M. S. (2021). "Semantic segmentation of tumors in kidneys using attention U-Net models," in *2021 5th International Conference on Electrical, Electronics, Communication, Computer Technologies and Optimization Techniques, ICEECCOT 2021 - Proceedings* 286–290. doi: 10.1109/ICEECCOT52851.2021.9708025
- Guo, J., Zeng, W., Yu, S., and Xiao, J. (2021). "RAU-Net: U-Net model based on residual and attention for kidney and kidney tumor segmentation," in *2021 IEEE International Conference on Consumer Electronics and Computer Engineering (ICCECE)* (IEEE) 353–356. doi: 10.1109/ICCECE51280.2021.9342530
- Gurcan, M. N., Boucheron, L. E., Can, A., Madabhushi, A., Rajpoot, N. M., and Yener, B. (2009). Histopathological image analysis: a review. *IEEE Rev. Biomed. Eng.* 2, 147–171. doi: 10.1109/RBME.2009.2034865
- Hatamizadeh, A., Terzopoulos, D., and Myronenko, A. (2020). Edge-Gated CNNs for volumetric semantic segmentation of medical images. *arXiv preprint arXiv:2002.04207*.
- Hatipoglu, N., and Bilgin, G. (2017). Cell segmentation in histopathological images with deep learning algorithms by utilizing spatial relationships. *Med. Biol. Eng. Comput.* 55, 1829–1848. doi: 10.1007/s11517-017-1630-1
- Heller, N., Sathianathan, N., Kalapara, A., Walczak, E., Moore, K., Kaluzniak, H., et al. (2019). The kits19 challenge data: 300 kidney tumor cases with clinical context, ct semantic segmentations, and surgical outcomes. *arXiv preprint arXiv:1904.00445*.
- Hou, X., Xie, C., Li, F., Wang, J., Lv, C., Xie, G., et al. (2020). "A triple-stage self-guided network for kidney tumor segmentation," in *2020 IEEE 17th International Symposium on Biomedical Imaging (ISBI)* (IEEE) 341–344. doi: 10.1109/ISBI45749.2020.9098609
- Hsiao, C.-H., Lin, P.-C., Chung, L.-A., Lin, F. Y.-S., Yang, F.-J., Yang, S.-Y., et al. (2022). A deep learning-based precision and automatic kidney segmentation system using efficient feature pyramid networks in computed tomography images. *Comput. Methods Progr. Biomed.* 221, 106854. doi: 10.1016/j.cmpb.2022.106854
- Isensee, F., and Maier-Hein, K. H. (2019). An attempt at beating the 3D U-Net. *arXiv preprint arXiv:1908.02182*.
- Kandel, I., and Castelli, M. (2020). How deeply to fine-tune a convolutional neural network: a case study using a histopathology dataset. *Appl. Sci.* 10, 3359. doi: 10.3390/app10103359
- Khan, A., Sohail, A., Zahoor, U., and Qureshi, A. S. (2020). A survey of the recent architectures of deep convolutional neural networks. *Artif. Intell. Rev.* 53, 5455–5516. doi: 10.1007/s10462-020-09825-6
- Kim, K. G. (2019). Deep learning book review. *Nature* 29, 1–73. doi: 10.4258/hir.2016.22.4.351
- Lv, Y., and Wang, J. (2021). "Three uses of one neural network: automatic segmentation of kidney tumor and cysts based on 3D U-Net," in *International Challenge on Kidney and Kidney Tumor Segmentation* (Cham: Springer International Publishing) 40–45. doi: 10.1007/978-3-030-98385-7_6
- Mayer, K. H., Safren, S. A., Elsesser, S. A., Psaros, C., Tinsley, J. P., Marzinke, M., et al. (2017). Optimizing pre-exposure antiretroviral prophylaxis adherence in men who have sex with men: results of a pilot randomized controlled trial of "Life-Steps for PrEP". *AIDS Behav.* 21, 1350–1360. doi: 10.1007/s10461-016-1606-4
- Mu, G., Lin, Z., Han, M., Yao, G., and Gao, Y. (2019). *Segmentation of Kidney Tumor by Multi-Resolution VB-Nets*. University of Minnesota Libraries Publishing. doi: 10.24926/548719.003
- Munien, C., and Viriri, S. (2021). Classification of hematoxylin and eosin-stained breast cancer histology microscopy images using transfer learning with EfficientNets. *Comput. Intell. Neurosci.* 2021, 5580914. doi: 10.1155/2021/5580914
- Myronenko, A., and Hatamizadeh, A. (2019). 3D kidneys and kidney tumor semantic segmentation using boundary-aware networks. *arXiv preprint arXiv:1909.06684*.
- Nazari, M., Jiménez-Franco, L. D., Schroeder, M., Kluge, A., Bronzel, M., and Kimiaei, S. (2021). Automated and robust organ segmentation for 3D-based internal dose calculation. *EJNMMI Res.* 11, 1–13. doi: 10.1186/s13550-021-00796-5
- Pandey, S., Singh, P. R., and Tian, J. (2020). Biomedical Signal Processing and Control An image augmentation approach using two-stage generative adversarial network for nuclei image segmentation. *Biomed. Signal Proc. Control* 57, 101782. doi: 10.1016/j.bspc.2019.101782
- Payer, C., Štern, D., Bischof, H., and Urschler, M. (2016). "Regressing heatmaps for multiple landmark localization using CNNs," in *International Conference on Medical Image Computing and Computer-Assisted Intervention* (Cham: Springer International Publishing) 230–238. doi: 10.1007/978-3-319-46723-8_27
- Rajnikanth, V., Vincent, P. M. D. R., Srinivasan, K., Prabhu, G. A., and Chang, C.-Y. (2023). A framework to distinguish healthy/cancer renal CT images using the fused deep features. *Front. Public Health* 11, 1109236. doi: 10.3389/fpubh.2023.1109236
- Ramadan, S. Z. (2020). Methods used in computer-aided diagnosis for breast cancer detection using mammograms: a review. *J. Healthcare Eng.* 2020, 9162464. doi: 10.1155/2020/9162464
- Ronneberger, O., Fischer, P., and Brox, T. (2015). "U-net: Convolutional networks for biomedical image segmentation," in *Medical Image Computing and Computer-Assisted Intervention—MICCAI 2015: 18th International Conference, Munich, Germany,*

- October 5-9, 2015, *Proceedings, Part III* 18 (Cham: Springer International Publishing) 234–241. doi: 10.1007/978-3-319-24574-4_28
- Sabarinathan, D., Parisa Beham, M., and Mansoor Roomi, S. M. M. (2020). Hyper vision net: kidney tumor segmentation using coordinate convolutional layer and attention unit. *Commun. Comput. Inform. Sci.* 1249, 609–618. doi: 10.1007/978-981-15-8697-2_57
- Santini, G., Moreau, N., and Rubeaux, M. (2019). Kidney tumor segmentation using an ensembling multi-stage deep learning approach. A contribution to the KiTS19 challenge. *arXiv preprint arXiv:1909.00735*. doi: 10.24926/548719.023
- Sharma, R., Halarnkar, P., and Choudhari, K. (2020). “Kidney and tumor segmentation using U-Net deep learning model,” in *5th International Conference on Next Generation Computing Technologies (NGCT-2019)*. doi: 10.2139/ssrn.3527410
- Tan, M., and Le, Q. (2019). “Efficientnet: Rethinking model scaling for convolutional neural networks,” in *International Conference on Machine Learning (PMLR)* 6105–6114.
- Tan, M., and Le, Q. V. (2021). “EfficientNetV2: smaller models and faster training,” in *International Conference on Machine Learning (PMLR)* 10096–10106.
- Thong, W., Kadoury, S., Piché, N., and Pal, C. J. (2018). Convolutional networks for kidney segmentation in contrast-enhanced CT scans. *Comput. Methods Biomech. Biomed. Eng.* 6, 277–282. doi: 10.1080/21681163.2016.1148636
- Torres, H. R., Queirós, S., Morais, P., Oliveira, B., Fonseca, J. C., and Vilaça, J. L. (2018). Kidney segmentation in ultrasound, magnetic resonance and computed tomography images: A systematic review. *Comput. Methods Programs Biomed.* 157, 49–67. doi: 10.1016/j.cmpb.2018.01.014
- Tsantekidis, A., Passalis, N., Tefas, A., Kannianen, J., Gabbouj, M., and Iosifidis, A. (2017). “Forecasting stock prices from the limit order book using convolutional neural networks,” in *2017 IEEE 19th Conference on Business Informatics (CBI)* (IEEE) 7–12. doi: 10.1109/CBI.2017.23
- Tsuneki, M. (2022). Deep learning models in medical image analysis. *J. Oral Biosci.* 64, 312–320. doi: 10.1016/j.job.2022.03.003
- Türk, F., Lüy, M., and Bar, işçi, N. (2020). Kidney and renal tumor segmentation using a hybrid v-net-based model. *Mathematics* 8, 1–17. doi: 10.3390/math8101772
- Üreyen, O., Dadal, i, E., Akdeniz, F, Sahin, T., Tekeli, M. T., Eliyatkin, N., et al. (2015). Co-existent breast and renal cancer. *Turkish J. Surg.* 31, 238–240. doi: 10.5152/UCD.2015.2874
- Valueva, M. V., Nagornov, N. N., Lyakhov, P. A., Valuev, G. V., and Chervyakov, N. I. (2020). Application of the residue number system to reduce hardware costs of the convolutional neural network implementation. *Mathem. Comput. Simul.* 177, 232–243. doi: 10.1016/j.matcom.2020.04.031
- Wen, J., Li, Z., Shen, Z., Zheng, Y., and Zheng, S. (2021). “Squeeze-and-excitation encoder-decoder network for kidney and kidney tumor segmentation in CT images,” in *International Challenge on Kidney and Kidney Tumor Segmentation* (Cham: Springer International Publishing) 71–79. doi: 10.1007/978-3-030-98385-7_10
- Xie, X., Li, L., Lian, S., Chen, S., and Luo, Z. (2020). SERU: A cascaded SE-ResNeXT U-Net for kidney and tumor segmentation. *Concurr. Comput.* 32, e5738. doi: 10.1002/cpe.5738
- Yang, G., Li, G., Pan, T., Kong, Y., Wu, J., Shu, H., et al. (2018). “Automatic segmentation of kidney and renal tumor in ct images based on 3d fully convolutional neural network with pyramid pooling module,” in *2018 24th International Conference on Pattern Recognition (ICPR)* (IEEE) 3790–3795. doi: 10.1109/ICPR.2018.8545143
- Zhang, Y., Wang, Y., Hou, F., Yang, J., Xiong, G., Tian, J., et al. (2019). Cascaded volumetric convolutional network for kidney tumor segmentation from CT volumes. *arXiv preprint arXiv:1910.02235*. doi: 10.24926/548719.004
- Zhao, W., Jiang, D., Queralta, J. P., and Westerlund, T. (2020). MSS U-Net: 3D segmentation of kidneys and tumors from CT images with a multi-scale supervised U-Net. *Inform. Med. Unlocked* 19, 100357. doi: 10.1016/j.imu.2020.100357