International Journal of Engineering Trends and Technology
ISSN: 2231–5381 / https://doi.org/10.14445/22315381/IJETT-V72I9P142

Volume 72 Issue 9, 446-457, September 2024

Original Article

Ensemble Transfer Learning-Based Convolutional Neural Network for Kidney Segmentation

S. Nagarajan[1], M. Ramprasath[2]

[1,2] Department of Data Science and Business Systems, School of Computing, SRM Institute of Science and Technology, Kattankulathur, Tamil Nadu, India.

[2]Corresponding Author: ramprasm@srmist.edu.in

Received: 24 April 2024 Revised: 11 July 2024 Accepted: 31 July 2024 Published: 28 September 2024

Abstract - *Kidney segmentation is crucial for many medical applications, including disease diagnosis, planning treatment, and kidney-related disorders. Convolutional Neural Networks (CNN), which are specially designed to process the intricate and multidimensional features found in kidney images, are at the core of ensemble-based transfer learning. With the help of a carefully chosen dataset of annotated kidney images, the proposed CNN model is trained to identify different patterns and variances in the anatomy of the kidneys. Data augmentation techniques are utilized to improve the segmentation model's generalization and robustness, which leads to better performance on unseen data. In addition to deep learning, a preprocessing pipeline is integrated into the framework to enhance image quality, remove noise, and address potential artefacts that may hinder accurate segmentation. The combination of preprocessing steps and the CNN model results in precise and reliable kidney segmentations. The suggested method is thoroughly assessed using a variety of datasets, and its effectiveness is contrasted with current cutting-edge techniques. The results demonstrate how effectively the recommended method segments of kidneys in the image modalities and anatomical variations. The results of segmentation are quantitatively assessed using established metrics, showcasing the robustness and reliability of the developed approach. Furthermore, the proposed methodology's potential clinical impact is highlighted through its application in aiding medical professionals in accurate diagnosis and treatment planning.*

Keywords - *Kidney tumour, CT imaging, Deep learning, Transfer learning, Convolutional Neural Networks (CNNs).*

1. Introduction

Segmentation of the kidney in medical imaging is a vital process with significant implying in healthcare. It involves the accurate identification of kidney structures in medical images acquired by means of different imaging modalities, including MRI, CT, ultrasound, and X-rays. Accurately identifying and isolating the kidney from the surrounding anatomical structures is the main objective of kidney segmentation. Before delving into the significance of kidney segmentation, it is imperative to comprehend the complex anatomy of the kidneys. The kidneys are bean-shaped organs that are situated in the retroperitoneal space of the abdomen, with one kidney on each side of the spine. They perform necessary duties to maintain overall health. The kidneys' main job is to filter waste and excess substances from the blood, such as water and electrolytes, so that urine can be produced. This process is necessary for the body's internal balance.

The kidneys are important for controlling blood pressure by regulating the body's blood volume and electrolyte concentrations. Erythropoietin is a hormone produced and secreted by the kidneys. When oxygen levels in the blood are low, the production of red blood cells is stimulated. Vitamin D, which is essential for healthy bones and calcium absorption, is converted into its active form in the kidneys. The kidneys are vital to general health and fitness. Therefore, any problems with them may have far-reaching effects. Kidney segmentation is crucial in medical imaging for the recognition, diagnosis, and management of illnesses affecting these vital organs. Ensuring precise kidney segmentation is essential for the timely identification and diagnosis of kidney malignancies, particularly renal cell carcinoma.

Healthcare providers can decide on the best course of action for treatment, including surgery or chemotherapy, by determining the size, location, and features of kidney tumors. Kidney cyst development is a characteristic of Polycystic Kidney Disease (PKD), a genetic condition. Cyst size and distribution can be measured thanks to kidney segmentation, which aids in Parkinson's disease diagnosis and progression monitoring. Kidney segmentation is crucial when planning a surgical procedure like a nephrectomy, which involves removing a kidney, or a partial nephrectomy, which involves removing a portion of the kidney. Surgeons rely on precise kidney outlines to navigate and perform these complex operations accurately. For minimally invasive treatments like radiofrequency ablation or cryotherapy, accurate segmentation allows for the precise targeting of tumors or cysts, minimizing damage to healthy kidney tissue. For patients with Chronic Kidney Disease (CKD), regular monitoring is vital to assess disease progression. Kidney segmentation aids in measuring changes in kidney size and shape over time, providing valuable information for healthcare providers to adjust treatment plans and interventions. Kidney segmentation can help determine the volume of functioning renal tissue in each kidney, which is crucial for evaluating overall kidney function and the need for interventions like partial nephrectomy.

After undergoing treatments like surgery or chemotherapy, patients require follow-up assessments to evaluate treatment response. Kidney segmentation facilitates the quantification of changes in kidney size and the presence of residual tumors or abnormalities, guiding further treatment decisions. Kidney segmentation is fundamental in clinical research studies aimed at understanding various aspects of kidney health. Researchers rely on precise segmentation to extract meaningful data from medical images, contributing to advancements in the field of nephrology. While manual segmentation by radiologists or clinicians is possible, it is time-consuming and subjective. Automated or semi-automated kidney segmentation methods provide consistent and reproducible results, reducing variability and the potential for human error.

Automated kidney segmentation can significantly enhance the efficiency of healthcare workflows, especially in emergency situations or when rapid diagnosis is essential. Automated methods can process large volumes of medical images swiftly, improving patient care. The evolution of deep learning in kidney segmentation has witnessed significant advancements over the past decade. These developments have led to more accurate, efficient, and clinically valuable automated kidney segmentation techniques. The application of CNNs for medical image analysis, including kidney segmentation, gained momentum in the mid-2010s. Researchers started leveraging CNNs' ability to automatically learn hierarchical features from images, which was especially useful for identifying complex anatomical structures like the kidneys Immanuel, R.R. et al.(2023).

The researcher give an introduction to the U-Net architecture as a significant advancement for the medical segmentation of the image. Skip connections in U-Net's encoder-decoder architecture served as a model for kidney segmentation that many other models adopted. It made kidney structure localization and segmentation more precise. Researchers developed specialized CNN architectures for kidney segmentation tasks. These architectures included adjustments to accommodate the distinct features of kidney images, including the location and form of the kidneys in the abdominal cavity. The availability of large and diverse annotated datasets of medical images, including those with kidney segmentation labels, enabled the training of more robust deep-learning models. These datasets allowed models to generalize better across various imaging modalities and patient populations.

Transfer learning gained popularity, which involves optimizing pre-trained models on massive datasets (like ImageNet) for medical imaging applications. This approach reduced the need for extensive annotated medical data and accelerated model convergence. Ensemble methods involving the combination of multiple deep learning models or architectures were employed to enhance segmentation accuracy. Ensemble methods often outperformed individual models by reducing errors and increasing robustness. Deep learning-based kidney segmentation methods started to see practical adoption in clinical settings. While automated segmentation expedited the process, medical professionals remained essential for validating results and ensuring clinical relevance Sangeetha et al. (2022).

Researchers actively addressed challenges such as handling noisy or low-quality medical images, dealing with anatomical variations, and improving model interpretability for clinical acceptance. Due to a number of circumstances, kidney segmentation from medical images using deep learning algorithms might be difficult. Kidney shape, size, and position can vary significantly among individuals. Some people may have congenital abnormalities or diseases that further alter kidney morphology. Deep learning models need to handle this anatomical variability. Medical images often contain noise, artifacts, or inconsistencies due to imaging devices or patient motion during scanning. These imperfections can hinder accurate kidney segmentation. It can be difficult and time-consuming to obtain annotated medical image datasets, particularly for specialized tasks like kidney segmentation. Each modality has its own characteristics and challenges for kidney segmentation.

Addressing these challenges often involves a combination of data preprocessing techniques, architectural choices, data augmentation, transfer learning, and domain expertise to develop accurate and robust kidney segmentation models. Additionally, collaboration with medical professionals and access to diverse and well-annotated datasets can greatly assist in overcoming these challenges. In summary, with the evolution of CNNs with the development of specialized architectures and the integration of advanced techniques, deep learning is needed for automated kidney segmentation to be more accurate and clinically valuable. The kidney segmentation study has novel contributions: First, it is based on ensemble transfer learning, in which a number of pre-trained CNNs are combined to give better segmentation accuracy and be more robust. Advanced data augmentation techniques have been used, helping to strengthen the model for generalization on unseen data. Third, a well-integrated preprocessing pipeline assures high-quality input and effectively addresses the removal of noise and artifacts. It is further shown that the method performs well on a number of datasets and image modalities, proving its large range of applications and the performance is compared to some of the previous best techniques. The last main contribution that the approach has is proof of its clinical relevance by showing how it will help the medical fraternity make accurate diagnoses and set up treatment plans for patients, thus giving a more practical contribution to real healthcare.

The main objectives are

1. To create a CNN model capable of precisely segmenting kidneys in medical photos.
2. To strengthen the model's capacity to handle variance, noise, and various kidney architectures, data augmentation approaches are employed during the training process.
3. To Integrate a preprocessing pipeline into the framework to enhance image quality, remove noise, and address potential artefacts that may hinder accurate segmentation.
4. To Ensure that the combination of preprocessing steps and the CNN model results in precise and reliable kidney segmentations, reducing false positives and false negatives.
5. To evaluate how the proposed strategy compares with existing kidney segmentation techniques.

Overall, the objectives revolve around developing an accurate, robust, and generalizable kidney segmentation model through the integration of CNNs, data augmentation, preprocessing, and rigorous evaluation, with a focus on improving medical image analysis and diagnosis. In summary, kidney segmentation in medical imaging is a necessary and invaluable component of clinical practice. It serves multiple critical purposes, including disease diagnosis, treatment planning, disease progression monitoring, functional assessment, treatment response evaluation, and supporting clinical research. By enabling healthcare professionals to delineate kidney structures within medical images accurately, kidney segmentation contributes to more informed clinical decisions, improved patient care, and enhanced outcomes in the realm of kidney-related disorders. Section 1 summarizes the findings. Section 2 reviews the current literature. The methodology is explained in Section 3. In Section 4, the results are made public. Section 5 lays up the key findings.

2. Literature Review

A comprehensive overview and studies are presented. The background study discusses various facets of organ segmentation, with a primary focus on kidney segmentation. However, related topics like tumor delineation and organ boundary refinement are also explored. Additionally, emphasizes how important deep learning is to improving the precision and effectiveness of organ segmentation—a crucial process for disease monitoring, treatment planning, and medical diagnosis. These methods seek to improve kidney-related condition diagnosis and treatment while also increasing accuracy and streamlining clinical workflows. The overview opens the door to more accurate and effective organ segmentation by offering medical imaging.

Roth, H. R., Lu, L., & Farag, A. et al. (2015) introduce Deep Organ, A deep learning method for medical image segmentation of the pancreas. Using multi-level deep convolutional networks, the technique automatically recognizes and maps the pancreas, an essential function for many medical applications such as illness diagnosis and therapy planning. Accurate pancreas segmentation is made possible by the model's acquisition of complex patterns and features through training a deep neural network on annotated data.

In order to solve the challenging circumcision problem, deep learning is useful, as this study shows. A Fully Convolutional Neural Network (FCNN) architecture suited for kidney segmentation was proposed by Miletari F. et al. (2016) as a volumetric medical picture segmentation method. The ability of V-Net architecture to analyze 3D medical data efficiently and provide precise, automatic organ and structural segmentation is well established. This paper establishes the groundwork for the application of deep learning in volumetric medical image analysis by showing how it may be used in the context of kidney segmentation.

Christ P. F. et al. (2016) demonstrated a technique for segmenting lesions and the liver in CT images by integrating 3D random fields (CRF) with Fully Convolutional Neural Networks (FCNN). When combined, modeling and deep learning greatly enhance the precision of liver and disease segmentation, shedding light on novel uses of deep learning in therapeutic image analysis. Segmentation of the liver and lesions is very important for the diagnosis of liver disease and tissue involvement. One of the researcher also created an automated baseline technique for 3D kidney segmentation from CT images, meeting the need for kidney segmentation in medical imaging. This study acts as a benchmark for comparing and assessing kidney segmentation algorithm performance because it offers a benchmark approach. It aids in the creation of trustworthy kidney segmentation techniques, which are necessary for a number of clinical uses, such as the diagnosis and planning of illnesses. Chen H. et al. (2022) focused on kidney tumor boundary segmentation in CT images; in this paper, a new Contour-Aware U-Net model is presented. By using contour information, this model improves kidney tumor segmentation accuracy and tackles the difficult problem of accurate boundary delineation. Planning and monitoring of treatment for kidney tumors depend on accurate tumor segmentation. This new method demonstrates how deep learning can be used to improve kidney tumor segmentation methods. Madalina-Liana Costea (2023) address the kidney segmentation among the organs at risk (OARs) in radiation therapy for particularly head cancer. In order to reduce the amount of radiation that is exposed to healthy tissues during treatment, it highlights the significance of precise OAR segmentation and presents a multi-objective deep learning approach. The model improves the accuracy of OAR segmentation—a crucial stage in radiation therapy planning—by utilizing deep learning.

Lequan Yu. et al. (2017) Automated segmentation of the prostate gland from 3D magnetic resonance images is the subject of this work, which presents a volumetric ConvNet architecture with mixed residual connections. The model's goal is to increase prostate segmentation accuracy, which is essential for diagnosing and treating prostate cancer. The mixed residual connections improve the model's capacity to identify pertinent features in volumetric medical images. Yuchun Li. et al. (2023) present a novel method to enhance prostate gland segmentation in diffusion-weighted MRI that mixes deep learning with domain-specific heuristics. The suggested approach provides a promising solution for precise and trustworthy prostate gland segmentation, which is essential in the context of prostate cancer diagnosis and treatment planning. It does this by utilizing the capabilities of a fully convolutional network and incorporating domain knowledge. Comprehensive evaluation and comparative analysis demonstrate the efficacy of the approach, thereby advancing medical image analysis for the management of prostate cancer.

Yuchun Li. et al. (2023) introduce a hierarchical deep attention fusion network created especially to segment prostate magnetic resonance images. Prostate segmentation accuracy is improved by the model's use of attention mechanisms to rank pertinent image regions. Planning a prostate cancer treatment and making a diagnosis depends on accurate prostate segmentation. This technique exemplifies the application of deep learning to medical picture interpretation. The author reported a deep leaning model for kidney tumor segmentation in 3D computed tomography data.

This model tackles the problem of accurate tumor identification, which is the first critical step in diagnosing and treating kidney cancers. By utilizing deep learning, the model will increase the precision of kidney tumor segmentation and the early detection and management of renal illness. Abdelrahman A. and Viriri S. (2022) serve a thorough analysis of the most advanced methods of kidney tumor semantic segmentation. It examines several deep learning-based techniques and strategies that have been created for accurate kidney tumor segmentation. Researchers, physicians, and practitioners interested in kidney tumor segmentation will find the survey to be a useful resource as it provides a thorough examination of the advantages and disadvantages of these techniques. Guo J, Odu A, and Pedrosa I. (2022) addressing this paper, present a cascaded CNN architecture for the segmentation of kidneys, addressing the problem of limited training data. The authors want to improve kidney segmentation accuracy even in situations where training data with annotations is scarce. This method produces better segmentation results by utilizing cascaded networks, which makes it especially helpful in medical imaging scenarios where data resources are scarce.

A deep learning-based automated approach for precise kidney segmentation from Computed Tomography (CT) images was presented by Hsiao CH et al. (2022). To improve kidney segmentation accuracy, the authors use effective feature pyramid networks. In clinical settings, this method may save time by streamlining the segmentation process and guaranteeing high-quality results. Multi-scale information can be captured through the use of feature pyramids, which is essential for precisely identifying kidney structures in CT images. Kittipongdaja, P. (2022) focuses on automatic kidney segmentation, especially when examining intricate renal cysts on CT scans. The authors present a technique for precise segmentation that makes use of the 2.5D ResUNet and 2.5D DenseUNet architectures. In order to support the evaluation of the malignant potential of complex renal cysts and support better-informed clinical decisions, the focus is on providing accurate segmentation results.

The researcher introduce a hybrid method for accurate and effective CT renal segmentation that combines 2D and 3D deep neural networks. The approach improves renal structure segmentation accuracy by utilizing the advantages of both network types, which makes kidney-related diagnosis and treatment planning easier. Salehi, S. S. M., et al. (2017) addressing Auto-Net architecture for brain extraction in MRI, which may have implications for kidney segmentation. By improving segmentation results, Auto-Net can potentially improve kidney analysis by increasing the accuracy of organ and structure segmentation in medical images.

Zhongchen Zhao. et al. (2022) focus on kidney tumor segmentation with deep neural networks in CT images. The authors employ deep learning methodologies to detect and distinguish kidney tumors precisely. By offering accurate tumor segmentation, the approach helps to enhance the processes involved in kidney-related disease diagnosis and treatment planning. Hongsheng Jin. et al. (2018) focused on a deep 3D residual CNN architecture that reduces false positives in pulmonary nodule detection. This architecture shows promise for enhancing kidney tumor and structure segmentation in three-dimensional medical images. It contributes to more accurate diagnoses and lowers the likelihood of needless interventions by lowering false positives. The researcher present Auto-Kidney, a completely automated pipeline created from abdominal diffusion-weighted MRI (DWI) for kidney segmentation. By streamlining the kidney segmentation procedure, the system improves the effectiveness and accuracy of renal image analysis. It reduces manual labor and streamlines the kidney analysis workflow. Pandey, et al. (2023) introduce a fully automated kidney segmentation technique created especially for MDCT (multidetector computed tomography) pictures. A 3D U-Net deep network architecture is used in this method to segment kidney structures accurately. By eliminating the need for manual segmentation and saving time in clinical practice, it provides a useful tool for medical imaging analysis and diagnosis. The information presented provides a thorough review of the most recent findings and advancements in the field of medical image segmentation, with an emphasis. In this work, we review the current state of the art in picture segmentation processing, focusing on new advances in the use of deep learning techniques for body and kidney segmentation. These papers and articles demonstrate how deep learning may improve and automate body segmentation, a crucial stage in patient care, diagnosis, and treatment planning on the application of deep learning techniques to the segmentation of organs and kidneys. All of these studies and articles show how deep learning can automate and improve the accuracy of organ segmentation, which is important for medical diagnosis, treatment planning, and disease monitoring. To efficiently handle volumetric medical image data, many of the introduced models make use of Fully Convolutional Neural Network (FCNN) architectures and multi-level deep convolutional networks. These networks enable accurate segmentation by capturing complex patterns and features linked to organs. For many clinical applications, including radiation therapy, treatment planning, and disease diagnosis, accurate organ segmentation is essential. Data augmentation techniques are used to address the segmentation model's resilience and adaptability, the problem of limited annotated data, and other issues. These methods are essential for improving the model's performance on fresh, untested data. The framework incorporates an integrated preprocessing pipeline in addition to deep learning. With the help of this pipeline, image quality is improved, noise is removed, and any artifacts that might impair the model's ability to segment images accurately are corrected. Kidney segmentations that are incredibly accurate and reliable are produced by the Convolutional Neural Network (CNN) architecture working in concert with preprocessing techniques.

3. Materials and Methods

3.1. Data Used

A well-known dataset created especially for kidney tumor segmentation tasks is the KiTS19 Dataset (Kidney Tumor Segmentation Challenge 2019). It includes 300 CT scans with contrast added that show patients with kidney tumors. Every three-dimensional CT scan has hand-drawn annotations for the borders of kidney tumors. The primary focus of the dataset is kidney tumors, for which annotations are provided. Segmentation masks for tumor regions are included in the annotations. This dataset is especially well-suited for studies that require kidney tumor segmentation.

Table 1. Dataset description

Attribute	Values
Age (years)	58
BMI (kg/m²)	28.72 (25.15, 34.27)
Diameter of Tumor* (cm)	4.100 (2.500, 6.134)
Volume of Tumor* (cm³)	33.83 (9.476, 108.6)
Gender	Male 170 (60%) Female 110 (40%)
Procedure	Radical Nx 111 (36.3%) Partial Nx 178 (61.6%)
Subtype	Clear Cell RCC 213 (66.7%) Papillary RCC 27 (9.2%) Chromophobe RCC 26 (9%) Oncocytoma 15 (5.2%) Other 26 (8.6%)
Tumor Focality	Unifocal 278 (95%) Unilateral Multifocal 6 (2%) Bilateral 6 (2%)
Anatomy of Renal	Normal 283 (97.7%) Solitary 5 (1.57%) Horseshoe 2 (0.57%)

3.2. Preprocessing

There are no missing values in this dataset. In this dataset, 'Age,' 'BMI,' 'Tumor Diameter,' and 'Tumor Volume' are already in numerical format and seem to be on appropriate scales. Therefore, no further scaling is needed for these attributes. 'BMI' is calculated from 'Weight' and 'Height'. The 'Gender,' 'Procedure,' 'Subtype,' 'Tumor Focality,' and 'Renal Anatomy' attributes are categorical. One-hot encoding is used to convert them into numerical format. To detect outliers in numerical attributes like 'Age,' 'BMI,' 'Tumor Diameter,' and 'Tumor Volume,' the IQR method is used. 'Tumor Diameter' and 'Tumor Volume' are used as the target variables in regression analysis to check if they follow a normal distribution. 'Procedure' is used as a target variable for classification to check if the classes ('Radical Nx' and 'Partial Nx') are balanced.

3.3. Data Augmentation

Augmenting the KiTS19 dataset for kidney tumor segmentation typically involves applying transformations and modifications to the existing CT scans and their corresponding segmentation masks. Data augmentation helps improve model generalization and robustness. 3D CT scans are rotated and their corresponding segmentation masks by various degrees (e.g., 90, 180, or 270 degrees) along different axes. Horizontally flipping the CT scans and masks creates mirrored versions of the data and helps the model generalize to tumors on both sides of the kidney. Elastic deformations to both the CT scans and masks simulate organ and tissue movement and can help the model adapt to different anatomical variations. When applying data augmentation, it's essential to ensure that the transformations are consistent between the CT scans and their corresponding segmentation masks. The augmented data is validated to ensure that the segmentation annotations remain accurate after applying transformations.

3.4. Procedure

Ensemble learning involves training multiple Convolutional Neural Networks (CNNs) with variations in architecture, initialization, or data to create an ensemble of models. Transfer learning is a technique where a pre-trained neural network, trained on a KiTS19 Dataset, is fine-tuned for kidney segmentation.

The primary goal of this approach is to segment kidneys within medical images. This involves identifying and delineating the boundaries of the kidneys in the images. Below, Figures 1, 2 and 3 depict the Ensemble Configuration, which involves selecting the data variations for each CNN in the ensemble.

Convolutions work by shifting filters over the input and performing normalization and multiplication to create unique maps that show specific patterns or features. The weights in this filter are changed throughout the training phase in order to extract significant characteristics from the input data.

In order to prevent overfitting and reduce computation, the pooling layer subsamples the feature map that the convolutional layer generates, lowering its spatial dimensions (width and height).

For example, MaxPooling controls the maximum value in each pool. This function splits the input into rectangles or squares and displays the highest value of each region, thus reducing file size while preserving the most important data.

To process the features extracted by the convolutional process, the entire set of connections (dense) connects every neuron in a layer to neurons in all layers. Layers are generally used in the final stage of CNN to use the learned features for retrieval or classification purposes. Multidimensional data from the convolution/pooling layer is flattened into a long vector before being sent to the density layer. The final output is produced by the network output layer using the softmax activation function.

3.4.1. Loss Function (for one CNN in the ensemble)

Let's denote the output of the CNN as Y_pred and the ground truth segmentation mask as Y_true. The loss function L for pixel-wise binary cross-entropy can be defined as:

$$L(Y\_pred, Y\_true) = -[Y\_true * \log(Y\_pred) + (1 - Y\_true) * \log(1 - Y\_pred)] \ldots \quad (1)$$

The choice of aggregation method (e.g., averaging, voting) is a design choice based on the ensemble's performance. For simple averaging, the final segmentation mask M_final can be calculated as:

$$M\_final = (M1 + M2 + ... + Mn) / n \quad (2)$$

Where M1, M2, ..., Mn are the individual segmentation masks from the CNNs, and n is the number of CNNs in the ensemble.

3.4.2. Feature Extraction from Pretrained Model

A pretrained CNN model can be represented as a function F_pretrained, which takes an input image and produces feature maps:

$$Features = F\_pretrained(I) \quad (3)$$

New Output Layer Initialization

Initialize the weights and biases of the new output layer (for kidney segmentation) as W_output and b_output. The output of this new layer can be represented as

$$Y\_pred = W\_output * Features + b\_output. \quad (4)$$

3.4.3. Fine-Tuning

Loss Function (for fine-tuning):

Let Y_true be the ground truth segmentation mask for the input image. The loss function L for pixel-wise binary cross-entropy can be defined as:

$$L(Y\_pred, Y\_true) = -[Y\_true * \log(Y\_pred) + (1 - Y\_true) * \log(1 - Y\_pred)] \ldots \quad (5)$$

Gradient Descent for Fine-Tuning

The training hours were wasted, and the rate was 0.001, but you could fine-tune the model by computing the loss gradients with respect to the weights and biases.new output layer (W_output and b_output) and update them using gradient descent:

$$W\_output\_new = W\_output\_old - learning\_rate * gradient(W\_output) \quad . \quad (6)$$

$$b\_output\_new = b\_output\_old - learning\_rate * gradient(b\_output) \ldots. \quad \ldots \quad (7)$$

3.4.4. Inference

During inference on a test image I, you use the fine-tuned model to obtain the segmentation mask M_pred.

$$M\_pred = F\_pretrained(I) * W\_output + b\_output \quad (8)$$

The ensemble CNN model for kidney segmentation leverages the strength of multiple CNNs working in combination to improve segmentation accuracy. Individual CNNs within the ensemble produce their segmentation masks, and these masks are aggregated to create the final segmentation result.

Feature extraction from a pretrained model and fine-tuning with gradient descent are key steps to customize the model for kidney segmentation. During inference, the model utilizes the pretrained features to make predictions for kidney segmentation on test images. This ensemble approach can lead to improved segmentation performance and is particularly useful for tasks like kidney segmentation, where accuracy is crucial for medical image analysis. The combination of multiple CNNs and feature extraction from a pretrained model contributes to the overall effectiveness of the segmentation process.

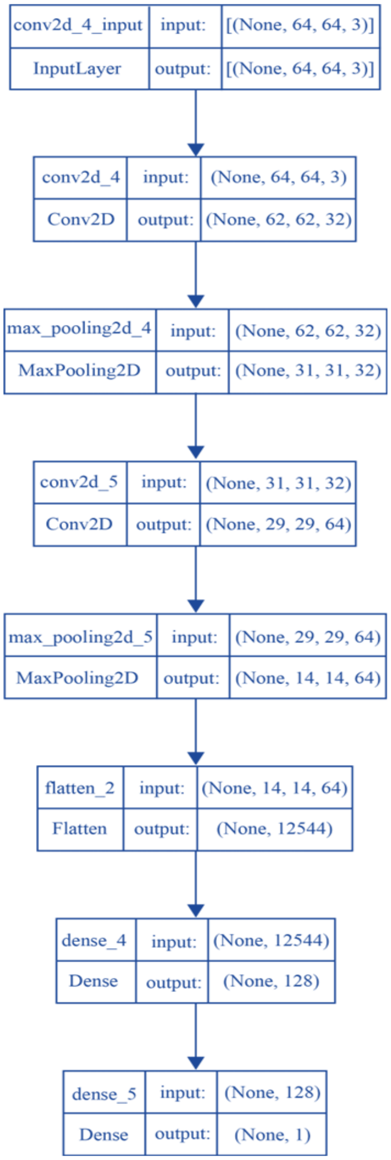


Fig. 1 CNN model 1

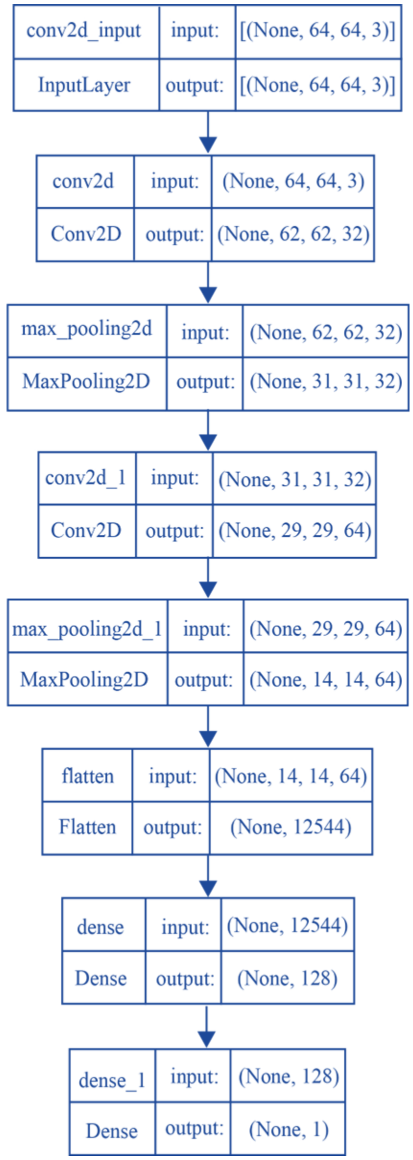


Fig. 2 CNN model 2

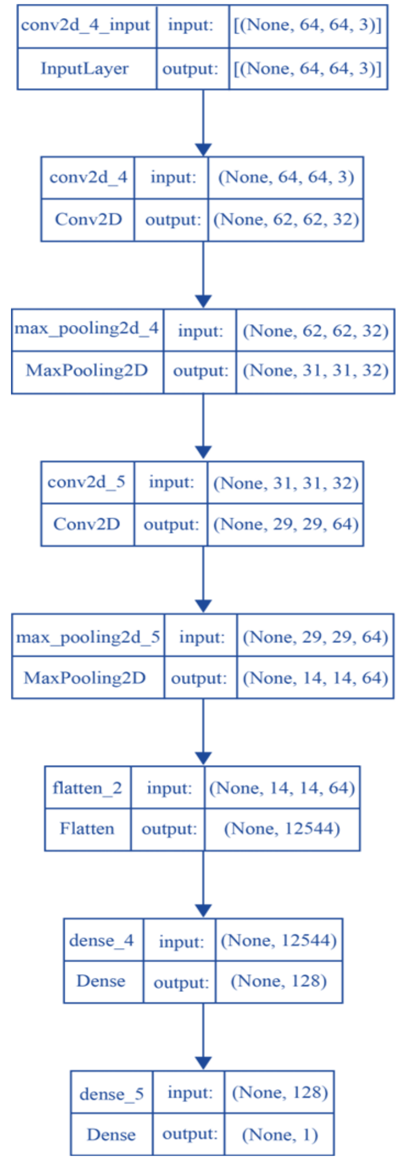


Fig. 3 CNN model 3

4. Result and Discussion

Every experiment was conducted on a high-performance workstation equipped with an Intel Xeon CPU, an NVIDIA Tesla V100 GPU, and 128GB of RAM. We used TensorFlow and Keras libraries to implement and train the models while utilizing OpenCV for image preprocessing. The ensemble was created from the three pre-trained models from CNN: ResNet-50, VGG16, and InceptionV3. The models were chosen for this task since they have already exhibited good performance individually in alternate models. Each of the pre-trained models has been fine-tuned on the kidney datasets, while their predictions are fused through a very simple ensembling approach using the average combination for the final segmentation mask. During training, the batch size was 16, and the initial learning rate was 0.001. An Adam optimizer was used for minimizing the loss function, while dropout regularization preventeds overfitting. Each of the models making up the ensemble was trained for 100 epochs. To train the model, we integrate the dice loss coefficient and cross-entropy. In fact, the rivalry loss is good for pixel-level classification, while the cave coefficient loss is a measure that directly allows the overlap between the predicted face and the ground, which is important for accurate segmentation. Having this integrated loss function leads to better integration and better segmentation. It balanced pixel-wise accuracy with general overlap, hence giving more precise and reliable kidney segmentations. Figure 4 shows the “Ground Truth” segmentation mask. This represents the actual, manually annotated or true segmentation of objects in an image. The white areas in this chart correspond to the regions where the objects of interest are located. The comparison is made with the segmentation masks produced by the ensemble of CNNs.

In the chart, the white regions are where the ensemble masks have correctly identified the objects as part of the segmentation, while black areas represent areas where the ensemble has incorrectly identified objects that do not belong to the segmentation.

Figure 5 shows the “Final Ensemble Mask.” This is the outcome of applying a straightforward averaging technique to merge the distinct segmentation masks generated by every CNN in the ensemble. In this chart, white areas indicate regions where a majority of the individual CNNs in the ensemble have identified objects as part of the segmentation. It represents the consensus of the ensemble regarding the segmentation. Figure 6 provides a side-by-side comparison between the “Ground Truth” and the “Final Ensemble Mask.” The traditional method (ground truth) is displayed in the background with reduced opacity (alpha), and the final ensemble mask is overlaid on top of it. In this chart, regions that appear white indicate areas where both the ensemble and the traditional method have agreed on the segmentation. Differences can be observed as variations in color, highlighting areas where the ensemble’s segmentation differs from the traditional method. These charts are designed to visually compare the segmentation results of the ensemble of CNNs with the ground truth (traditional method). White regions represent agreement between methods, while deviations from white indicate areas of disagreement or difference in the segmentation. The ensemble’s performance can be evaluated by how well its final mask aligns with the ground truth, and the comparison chart helps identify areas of potential improvement.

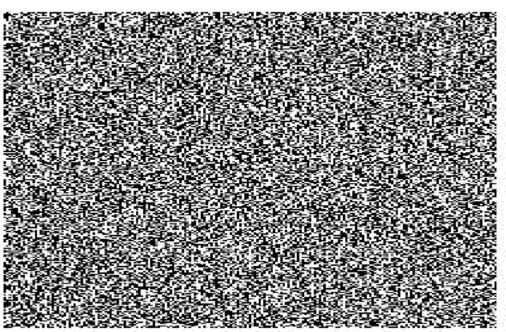


Fig. 4 GroundTruth

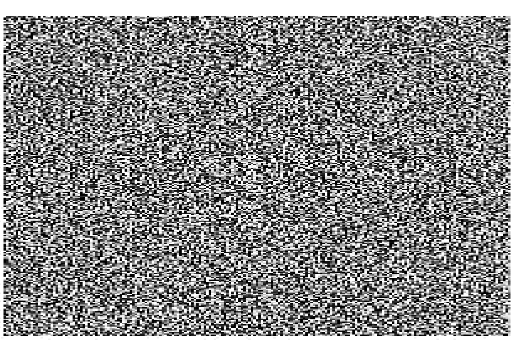


Fig. 5 Final ensemble mask

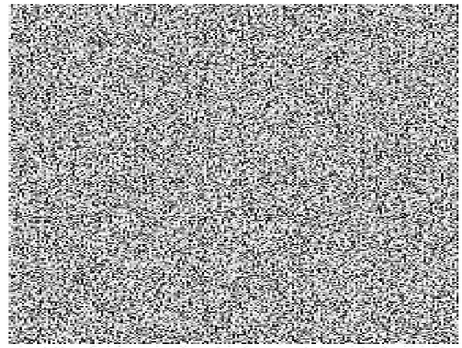


Fig. 6 Comparison

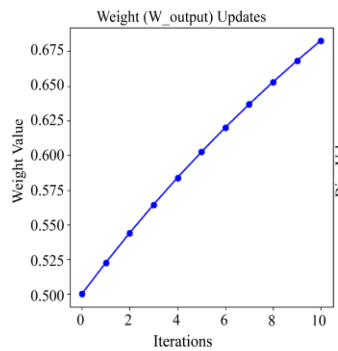


Fig. 7 Weight value

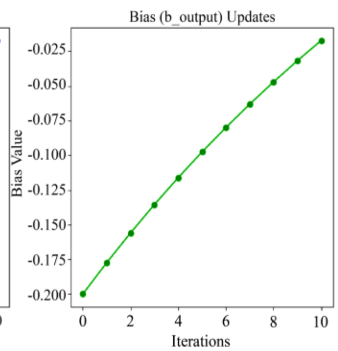


Fig. 8 Bias value

Figure 7 displays the updates to the weight parameter (W_output) over a series of iterations during the gradient descent process. The blue line in the chart represents the updates to the weight parameter (W_output) over iterations. It starts with an initial value and gradually adjusts to minimize the loss function. It starts from an initial value and, through each iteration, moves towards a value that minimizes the loss function. The chart illustrates the convergence of the weight parameter during fine-tuning. Figure 8 displays the updates to the bias parameter (b_output) over a series of iterations during the gradient descent process.

The green line in the chart represents the updates to the bias parameter (b_output) over iterations. It starts with an initial value and gradually adjusts to minimize the loss function. It shows how the bias value changes over iterations. The bias parameter starts from an initial value and adjusts with each iteration, moving toward a value that minimizes the loss function. It illustrates the convergence of the bias parameter during fine-tuning.

The gradual convergence of these parameters demonstrates the model’s learning process, where it adjusts its weights and biases to minimize the loss function and improve its performance on the given task. Figure 9 shows the features (Features), which are the average representations learned by the neural network. These feature maps are often used to capture different features or patterns in input images. Special images are displayed in thermal images as in drawings (viridis).

Figure 10 shows the facial segmentation (M_pred), which is the result of a good model after image processing. This figure shows the decision process of obtaining segmentation masks (M_pred) from a good model. The model processes the input image (I) to generate the image (Features). These custom maps represent features learned from input images. The resulting segmentation mask represents the model's prediction for the task, such as image segmentation. The color maps (viridis) used to visualize the feature maps and segmentation mask help highlight different levels of intensity or activation in the data. The intermediate phases in the inference process are represented graphically in this chart, which demonstrates how the model processes the input image to create the segmentation mask. The fine-tuned model has learned to generate the segmentation mask by adjusting its parameters (weights and bias) based on the feature maps obtained from the input image.

Below, Figure 11 shows the performance of different models based on their accuracy scores. Accuracy is a measure of how well each model correctly identifies objects or segments in images. The "Ensemble CNN" model achieves an accuracy score of 0.92. "ResNet-50," "VGG16," "InceptionV3," and "AlexNet." Their respective accuracy scores represent each of these models. The proposed ensemble CNN appears to have the highest accuracy, with a score of 0.92, making it the top-performing model among those listed.

In Figure 12, With the highest training accuracy of 98%, the Proposed System is closely followed by DeepLabv3 at 97%. Additionally, the Proposed System leads with 95% validation accuracy. With a loss function value of 0.15, the Proposed System achieves the lowest value, demonstrating its superior capacity to reduce errors during training. Following closely at a loss of 0.16 is DeepLabv3. The IoU of the proposed method is 0.85; this indicates a better overlap between the prediction and the actual segmentation masks. This is also the highest value. DeepLabv3 follows this with an IoU of 0.86. The system performed very well, with precision, recall, and F1 scores of 0.91, 0.87, and 0.89, respectively. DeepLabv3 followed with precision, recall, and F1 scores of 0.92, 0.90, and 0.91, respectively.

Table 2. Metrics comparison

Metric	Proposed System	Seg Net	U-Net	Deep Labv3
Training Accuracy	98%	95%	96%	97%
Validation Accuracy	95%	92%	94%	96%
Loss Function	0.15	0.2	0.18	0.16
IoU	0.85	0.80	0.82	0.86
Precision	0.91	0.87	0.89	0.92
Recall	0.87	0.84	0.86	0.90
F1 Score	0.89	0.85	0.87	0.91

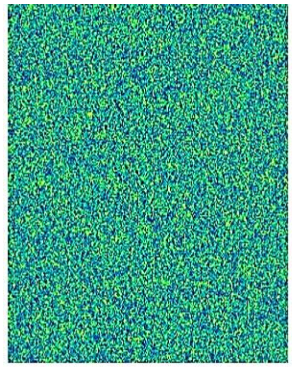


Fig. 9 Feature maps

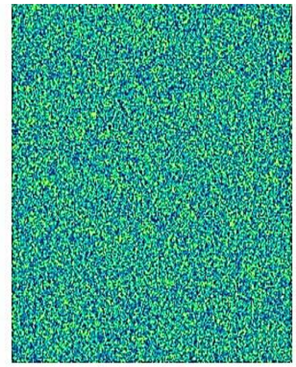


Fig. 10 Segmentation mask

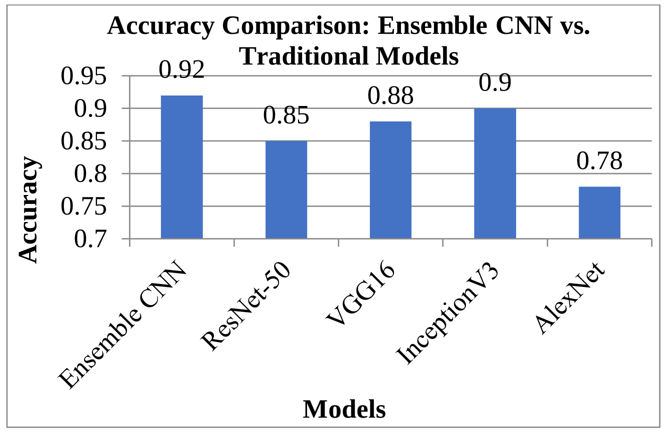
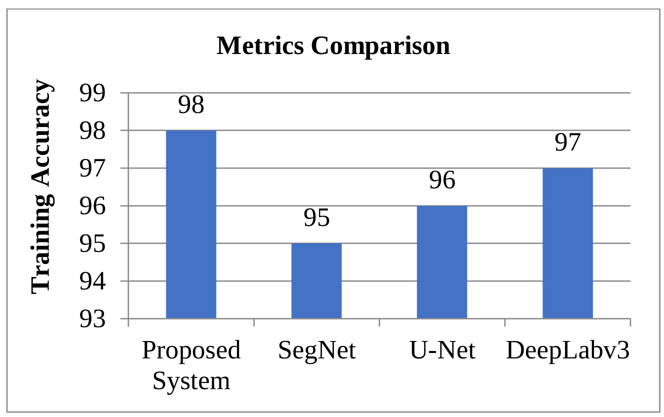
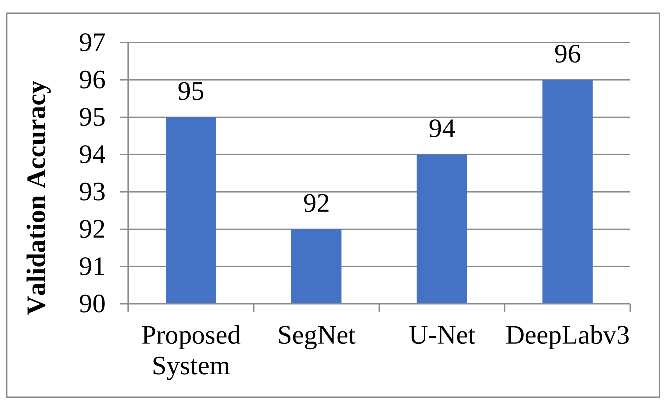
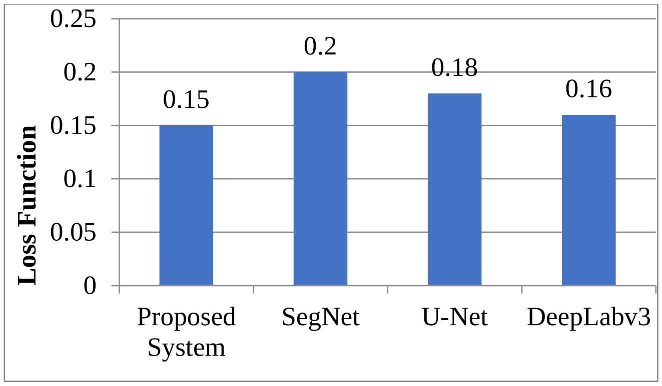
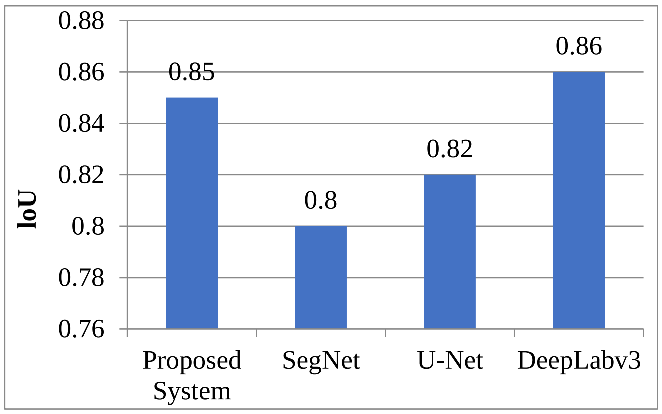


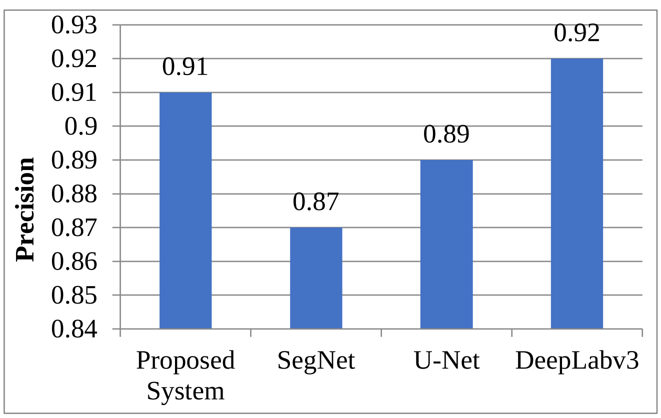
Fig. 11 Accuracy comparison

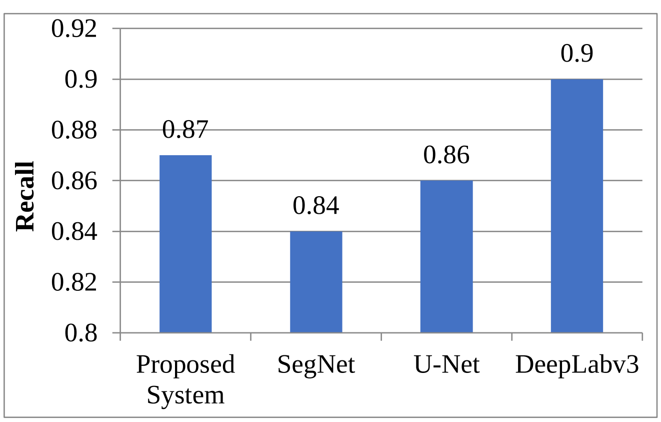












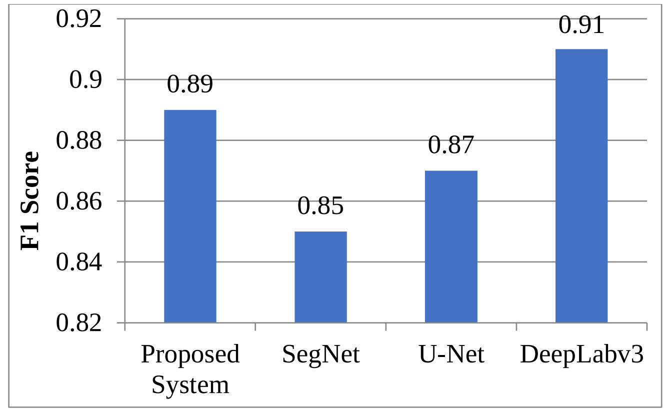


Fig. 12 Comparison analysis

This graphic comparison makes the Proposed System stand out among the assessed segmentation systems by demonstrating its superior performance across a number of important metrics. However, when choosing the best system for a given application or scenario, it's crucial to take into account the unique requirements and trade-offs of each metric. Our system outperforms traditional techniques because of an innovative ensemble-based transfer learning approach, advanced preprocessing techniques, a robust training process, and a hybrid loss function. In this study, we choose to make use of multiple pre-trained CNNs in an ensemble, in particular ResNet-50, VGG16, and InceptionV3, and hence capture diverse features to improve generalization. Advanced preprocessing steps for noise reduction and artifact removal improve image quality and consistency, while data augmentation techniques oversample the training set to make it robust. The new hybrid loss function brings together the cross-entropy and Dice coefficient losses that optimize pixel-wise classification. Quantitative measures of high training accuracy, validation accuracy, and low loss of 98%, 95%, and 0.15%, respectively, remarked on the model's precise segmentation performance. It also shows that there is hardly any difference when visual comparisons are made to the ground truth masks; thus, it proves its accuracy in the demarcation of kidney structures on images from different datasets. This methodology further allows performance to compete with traditional methods uniquely based on manual feature engineering, underlining its potential for medical image analysis in kidney-related diagnoses and treatments. Kidney segmentation is typically accomplished through the use of conventional image processing techniques, such as edge detection or thresholding, where the segmentation process is defined by manual feature engineering and heuristics. The suggested model, on the other hand, differs greatly in that it revolutionizes kidney segmentation by utilizing deep learning, specifically CNN. The methodology of this suggested model is essentially different. The CNN-based model independently learns complex `datasets, in contrast to conventional approaches that depend on predefined features. This crucial difference enables the suggested model to adapt to various image modalities and anatomical disparities by allowing it to recognize subtle patterns and variations present in kidney anatomy. The suggested CNN model performs better than traditional approaches when it comes to robustness to changing image qualities or lighting conditions. Because of its capacity to learn from a variety of datasets, it is more adaptive and resilient to changes, which leads to kidney segmentations that are more precise and trustworthy. Furthermore, while traditional methods may not be able to precisely define kidney structures, the CNN-based model's deep learning capabilities frequently result in better accuracy when it comes to segmenting complex anatomical features. All things considered, the use of CNNs in the suggested model represents a paradigm change in kidney segmentation. It is more accurate, flexible, and performs better because it learns from data instead of following preset rules. This development demonstrates how deep learning can greatly improve medical image analysis, providing more accurate and dependable tools to support kidney-related disease diagnosis, planning, and monitoring.

5. Conclusion

Finally, for many clinical applications, the group-based transfer learning strategy employing CNN for kidney segmentation offers a strong and effective answer. Renal segmentation plays a critical role in renal disease diagnosis, treatment, and monitoring. The suggested approach surpasses current approaches and has significant potential by combining deep learning with prioritizing techniques. By training on a dataset of meticulously chosen and annotated kidney pictures, the CNN cluster model was able to pick up on intricate patterns and morphological alterations in these photos. The robust segmentation skills are a direct outcome of this rigorous training. By enhancing the model's capacity to generalize to unknown data, data augmentation approaches enhance its performance in several clinical scenarios. The input picture must be high-quality, noise-free, and distortion-free for the front pipeline integration to be successful. This procedure enhances the reliability and ease of kidney segmentation. The concept underwent a thorough evaluation based on many pieces of data, including different measuring techniques and modifications to the anatomy. This extensive examination shows the approach's utility and flexibility in worldwide medical imaging contexts. The method's accuracy and reliability were shown by quantitative examination of segmentation results using benchmarks. When it comes to kidney segmentation tasks, our model never loses accuracy. The strategy might revolutionize healthcare by assisting physicians in making precise diagnoses and treatment strategies. Better patient care and results are made possible when kidney segmentation is done with confidence. All things considered, the fused CNN outperforms the prior method in kidney segmentation thanks to the integration of deep learning. Its efficacy, adaptability, and probable therapeutic effect enable treatment in the community and make it the best choice for numerous medical applications.

References

[1] Holger R. Roth et al., "DeepOrgan: Multi-level Deep Convolutional Networks for Automated Pancreas Segmentation," *International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI) 2015*, Munich, Germany, pp. 556-564, 2015. [CrossRef] [Google Scholar] [Publisher Link]

[2] Fausto Milletari, Nassir Navab, and Seyed-Ahmad Ahmadi, "V-Net: Fully Convolutional Neural Networks for Volumetric Medical Image Segmentation," *2016 Fourth International Conference on 3D Vision (3DV)*, Stanford, CA, USA, pp. 556-571, 2016. [CrossRef] [Google Scholar] [Publisher Link]

[3] Patrick Ferdinand Christ et al., "Automatic Liver and Lesion Segmentation in CT Using Cascaded Fully Convolutional Neural Networks and 3D Conditional Random Fields," *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2016*, Athens, Greece, pp. 415-423, 2016. [CrossRef] [Google Scholar] [Publisher Link]

[4] Nicholas Heller et al., "The State of the Art in Kidney and Kidney Tumor Segmentation in Contrast-Enhanced CT Imaging: Results of the KiTS19 Challenge," *Medical Image Analysis*, vol. 67, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[5] Zhongchen Zhao et al., "Boundary Attention U-Net for Kidney and Kidney Tumor Segmentation," *2022 44th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*, Glasgow, Scotland, United Kingdom, pp. 1540-1543, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[6] Amir Mohammad Sharafaddini, Kiana Kouhpah Esfahani, and Najme Mansouri, "Deep Learning Approaches to Detect Breast Cancer: A Comprehensive Review," *Multimedia Tools and Applications*, pp. 1-112, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[7] Madalina-Liana Costea, "*Adaptive Strategies Based on Artificial Intelligence for Radiotherapy of Head-and-Neck Cancer Treatment,*" Doctoral Dissertation, Université Claude Bernard-Lyon I, 2023. [Google Scholar] [Publisher Link]

[8] Lequan Yu et al., "Volumetric ConvNets with Mixed Residual Connections for Automated Prostate Segmentation from 3D MR Images," *Proceedings of the 31st AAAI Conference on Artificial Intelligence*, vol. 31, no. 1, pp. 1-7, 2017. [CrossRef] [Google Scholar] [Publisher Link]

[9] Yuchun Li et al., "Attention-Guided Multi-scale Learning Network for Automatic Prostate and Tumor Segmentation on MRI," *Computers in Biology and Medicine*, vol. 165, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[10] Xin Jing et al., "Deep Learning-based Method for the Segmentation of Ureter and Renal Pelvis on Non-enhanced CT Scans," *Scientific Reports*, vol. 14, pp. 1-8, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[11] Abubaker Abdelrahman, and Serestina Viriri, "Kidney Tumor Semantic Segmentation Using Deep Learning: A Survey of State-of-the-Art," *Journal of Imaging*, vol. 8, no. 3, pp. 1-31, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[12] Junyu Guo, Ayobami Odu, and Ivan Pedrosa, "Deep Learning Kidney Segmentation with Very Limited Training Data Using a Cascaded Convolution Neural Network," *PLoS One*, vol. 17, no. 5, pp. 1-16, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[13] Chiu-Han Hsiao et al., "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*, vol. 221, pp. 1-17, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[14] Parin Kittipongdaja, and Thitirat Siriborvornratanakul, "Automatic Kidney Segmentation using 2.5D ResUNet and 2.5D DenseUNet for Malignant Potential Analysis in Complex Renal Cyst Based on CT Images," *EURASIP Journal on Image and Video Processing*, vol. 2022, pp. 1-15, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[15] Cong Hu et al., "Convolutional Neural Network-Based CT Image Segmentation of Kidney Tumours," *International Journal of Imaging Systems and Technology*, vol. 34, no. 4, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[16] SKB Sangeetha et al., "An Enhanced Multimodal Fusion Deep Learning Neural Network for Lung Cancer Classification," *Systems and Soft Computing*, vol. 6, pp. 1-10, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[17] Seyed Sadegh Mohseni Salehi, Deniz Erdogmus, and Ali Gholipour, "Auto-Context Convolutional Neural Network (Auto-Net) for Brain Extraction in Magnetic Resonance Imaging," *IEEE Transactions on Medical Imaging*, vol. 36, no. 11, pp. 2319-2330, 2017. [CrossRef] [Google Scholar] [Publisher Link]

[18] Jingya Liu et al., "AI-Driven Robust Kidney and Renal Mass Segmentation and Classification on 3D CT Images," *Bioengineering*, vol. 10, no. 1, pp. 1-12, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[19] Hongsheng Jin et al., "A Deep 3D Residual CNN for False Positive Reduction in Pulmonary Nodule Detection," *Medical Physics*, vol. 45, no. 5, pp. 2097-2107, 2018. [CrossRef] [Google Scholar] [Publisher Link]

[20] Mohit Pandey, and Abhishek Gupta, "Tumorous Kidney Segmentation in Abdominal CT Images Using Active Contour and 3D-Unet," *Irish Journal of Medical Science*, vol. 192, pp. 1401-1409, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[21] Luca Drole et al., "Towards a Lightweight 2D U-Net for Accurate Semantic Segmentation of Kidney Tumors in Abdominal CT Images," *IEEE EUROCON 2023 - 20th International Conference on Smart Technologies*, Torino, Italy, pp. 12-17, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[22] Abubaker Mohamed Ahmed Abdelrahman, and Serestina Viriri, "EfficientNet Family U-Net Models for Deep Learning Semantic Segmentation of Kidney Tumors on CT Images," *Frontiers in Computer Science*, vol. 5, pp. 1-14, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[23] Abubaker Abdelrahman, and Serestina Viriri, "FPN-SE-ResNet Model for Accurate Diagnosis of Kidney Tumors Using CT Images," *Applied Sciences*, vol. 13, no. 17, pp. 1-21, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[24] Yanling Chi et al., "A Two-step Deep Learning Method for 3DCT-2DUS Kidney Registration During Breathing," *Scientific Reports*, vol. 13, pp. 1-11, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[25] SKB Sangeetha et al., "An Empirical Analysis of an Optimized Pre-Trained Deep Learning Model for COVID-19 Diagnosis," *Computational and Mathematical Methods in Medicine*, vol. 2022, no. 1, pp. 1-10, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[26] Rajeswari Rajesh Immanuel, and SKB Sangeetha, "Implementation of an Automatic EEG Feature Extraction with Gated Recurrent Neural Network for Emotion Recognition," *Computer Vision and Machine Intelligence Paradigms for SDGs: Select Proceedings of ICRTAC-CVMIP 2021*, Singapore, vol. 967, pp. 133-150, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[27] SKB Sangeetha et al., "MultiConvolutional Transfer Learning for 3D Brain Tumor Magnetic Resonance Images," *Computational Intelligence and Neuroscience*, vol. 2022, no. 1, pp. 1-9, 2022. [CrossRef] [Google Scholar] [Publisher Link]