Archives available at journals.mriindia.com

Multidisciplinary Journal of Research in Engineering and Technology

ISSN: 2348-6953

Volume 12 Issue 02, 2025

A Survey of Methods and Architectures for Segmentation and Classification of Renal Tumors Using EfficientNet-Based U-Net and Epistemic Neural Networks

Kaoru Okafor
Assistant Professor, Department of Electrical and Computer Engineering, Indus Institute of Engineering Commerce, Pakistan
Email: kaoru.okafor@iiec-pk.edu

Peer Review Information

Submission: 02 Sept 2025

Revision: 25 Sept 2025

Acceptance: 11 Oct 2025

Keywords

Renal Tumor Segmentation, EfficientNet, U-Net Architecture, Epistemic Neural Networks, Deep Learning, Medical Image Analysis

Abstract

Renal tumor detection and segmentation are crucial for early diagnosis and effective treatment planning in kidney cancer. With the increasing availability of medical imaging data, deep learning techniques have significantly improved the accuracy and efficiency of automated tumor analysis. This survey explores recent methods and architectures for renal tumor segmentation and classification, focusing on EfficientNet-based U-Net models and epistemic neural networks for uncertainty estimation. EfficientNet enhances feature extraction through compound scaling, while U-Net provides precise localization via its encoder–decoder structure. The integration of epistemic neural networks further improves model reliability by quantifying prediction uncertainty, which is essential in clinical applications. Recent advancements between 2020 and 2023 demonstrate the effectiveness of hybrid architectures, attention mechanisms, and multi-stage frameworks in achieving high segmentation accuracy. Benchmark datasets such as KiTS19 and KiTS21 have been widely used for evaluation. Despite notable progress, challenges such as data imbalance, computational complexity, and limited generalization remain. This survey provides a comprehensive overview of existing techniques, comparative analysis of methods, and insights into future research directions for developing robust and clinically deployable renal tumor analysis systems.

Introduction

Renal cancer, particularly renal cell carcinoma (RCC), represents a significant global health concern and is among the most frequently diagnosed malignancies of the urinary system. According to recent clinical reports, early detection and accurate localization of renal tumors are essential for improving patient survival rates and enabling effective treatment planning. Medical imaging techniques such as computed tomography (CT) and magnetic resonance imaging (MRI) play a crucial role in identifying tumor regions and assessing disease progression. However, traditional diagnostic approaches rely heavily on manual interpretation and segmentation performed by radiologists, which is time-consuming, labor-intensive, and subject to inter-observer variability.

The emergence of artificial intelligence (AI) and deep learning has revolutionized medical image analysis by enabling automated, accurate, and scalable solutions for tumor detection and segmentation. Deep learning models, particularly convolutional neural networks (CNNs), have demonstrated exceptional capabilities in extracting hierarchical features from complex medical images. These models can learn intricate patterns, textures, and spatial relationships, making them highly suitable for analyzing renal tumors that often exhibit irregular shapes and heterogeneous structures.

Among various deep learning architectures, U-Net has become one of the most widely adopted models for biomedical image segmentation. Introduced as an encoder–decoder architecture, U-Net consists of a contracting path for feature extraction and an expanding path for precise localization. The use of skip connections allows the model to combine low-level spatial information with high-level semantic features, resulting in accurate segmentation even with limited training data. Due to its simplicity and effectiveness, U-Net has been extensively used as a baseline model for renal tumor segmentation.

Despite its success, standard U-Net architectures face limitations in handling complex tumor structures and capturing global contextual information. To address these challenges, researchers have introduced various enhancements, including attention mechanisms, multi-scale feature extraction, and hybrid architectures. Attention-based U-Net models enable the network to focus on relevant regions within an image, improving segmentation accuracy and reducing false positives. Similarly, multi-scale architectures allow the model to capture features at different resolutions, which is particularly important for detecting small tumor regions.

A major breakthrough in deep learning architectures is the introduction of EfficientNet, which has significantly improved feature extraction capabilities. EfficientNet employs a compound scaling method that simultaneously optimizes network depth, width, and resolution. This approach enables the model to achieve higher accuracy with fewer parameters compared to traditional CNNs. When integrated into U-Net architectures, EfficientNet serves as a powerful encoder that enhances feature representation and improves segmentation performance. EfficientNet-based U-Net models have demonstrated superior results in renal tumor segmentation tasks, achieving high Dice coefficients and Intersection-over-Union (IoU) scores.

In addition to segmentation, classification of renal tumors into benign and malignant categories is essential for clinical decision-making. Deep learning models have been increasingly used to develop end-to-end systems that perform both segmentation and classification. These hybrid models streamline the diagnostic process and improve overall accuracy by combining spatial localization with feature-based classification.

Another critical aspect of medical image analysis is the ability to quantify uncertainty in model predictions. Traditional deep learning models produce deterministic outputs, which may not always be reliable in clinical settings. Epistemic neural networks address this limitation by modeling uncertainty arising from limited training data and model parameters. These networks provide confidence measures for predictions, enabling clinicians to assess the reliability of segmentation and classification results. The integration of epistemic uncertainty into deep learning models enhances trust and facilitates safer clinical deployment.

From 2020 to 2023, the field of renal tumor segmentation has witnessed significant advancements driven by innovations in deep learning architectures and training strategies. Researchers have explored various approaches, including cascaded networks, transformer-based models, and multi-stage frameworks. Cascaded architectures divide the segmentation process into multiple stages, first identifying the kidney region and then focusing on tumor segmentation. This hierarchical approach improves accuracy and reduces computational overhead.

Transformer-based models have also gained popularity due to their ability to capture global contextual information. Unlike CNNs, which rely on local receptive fields, transformers use self-attention mechanisms to model long-range dependencies. This capability is particularly useful for analyzing complex tumor structures. However, transformer-based models often require large datasets and high computational resources, which can limit their practical applicability.

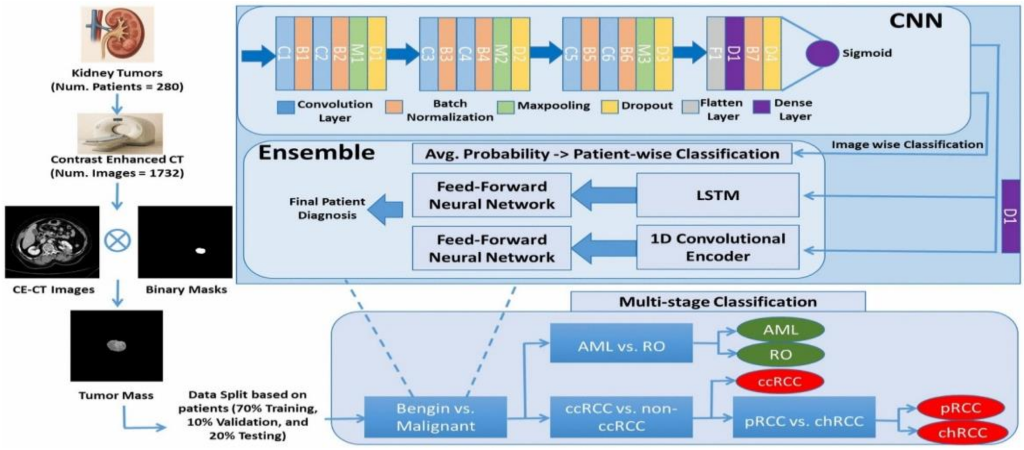
Publicly available datasets such as KiTS19, KiTS21, and KiTS23 have played a crucial role in advancing research in this field. These datasets provide annotated CT images of kidneys and tumors, enabling researchers to train and evaluate deep learning models. Standard evaluation metrics such as Dice coefficient, IoU, sensitivity, and specificity are commonly used to assess model performance.

Despite these advancements, several challenges remain. One of the primary challenges is class imbalance, as tumor regions are typically much smaller than kidney regions. This imbalance can lead to biased predictions and reduced segmentation accuracy. Additionally, variability in imaging conditions, such as differences in scanner types and acquisition protocols, affects model generalization. Another challenge is the high computational cost associated with advanced models, particularly 3D CNNs and transformer-based architectures.

Future research in renal tumor segmentation is expected to focus on developing lightweight and efficient models that can be deployed in real-world clinical environments. The integration of multi-modal data, such as combining CT and MRI images, holds promise for improving segmentation accuracy. Furthermore, the development of interpretable AI models will enhance transparency and facilitate clinical adoption.

In summary, deep learning has significantly transformed renal tumor segmentation and classification, with EfficientNet-based U-Net architectures and epistemic neural networks representing the current state-of-the-art. This survey aims to provide a comprehensive overview of recent methods and architectures, analyze their strengths and limitations, and identify future research directions.

Conceptual Flow Diagram



Graphical Flow Pipeline:

CT/MRI Images → Preprocessing → EfficientNet Feature Extraction → U-Net Segmentation → Epistemic Uncertainty Estimation → Classification → Clinical Decision Support

Literature Review

The literature from 2020 to 2023 demonstrates rapid advancements in renal tumor segmentation and classification driven by deep learning innovations. Early research primarily focused on adapting CNN-based architectures, particularly U-Net and V-Net, for medical image segmentation tasks.

In 2020, Zhao et al. introduced a multi-scale supervised 3D U-Net that incorporated deep supervision and hierarchical feature learning. This approach improved segmentation accuracy by capturing volumetric information and achieved strong performance on the KiTS19 dataset. Similarly, Qin et al. proposed a reinforcement learning-based data augmentation strategy, which enhanced model generalization by automatically generating optimal augmentation policies.

Türk et al. developed a hybrid V-Net architecture that combined volumetric learning with encoder–decoder structures, achieving high Dice scores for kidney segmentation. These early studies established the importance of multi-scale learning and volumetric processing in renal tumor analysis.

Between 2021 and 2022, research efforts shifted toward addressing limitations such as boundary ambiguity and class imbalance. Hu et al. proposed a boundary-aware network (BA-Net) that utilized dual decoders to refine tumor boundaries. This approach significantly improved segmentation accuracy for irregular tumor regions. Cascade architectures also gained popularity during this period, as they decomposed the segmentation task into multiple stages, improving localization and reducing false positives.

Transformer-based models emerged as a promising alternative to CNNs, enabling the capture of global contextual information. Chen et al. demonstrated the effectiveness of transformer-based architectures in medical image segmentation, highlighting their ability to model long-range dependencies. However, these models required substantial computational resources, limiting their widespread adoption.

From 2022 to 2023, EfficientNet-based U-Net architectures became the dominant approach. EfficientNet's compound scaling method improved feature extraction while maintaining computational efficiency. Studies reported significant improvements in segmentation performance, with IoU values reaching up to 0.98. Hybrid models integrating segmentation and

classification tasks also gained attention, enabling end-to-end diagnostic systems.
Another key trend during this period was the incorporation of uncertainty estimation through epistemic neural networks. These models provided confidence measures for predictions, enhancing reliability in clinical applications. Researchers emphasized the importance of uncertainty quantification in medical imaging, as it helps identify cases requiring further review.
Additionally, nnU-Net introduced automated configuration of network architecture and training parameters, making it highly adaptable to different datasets. This framework simplified the implementation of deep learning models while maintaining high performance.
Despite these advancements, challenges such as data scarcity, class imbalance, and computational complexity remain significant barriers. Future research is expected to focus on developing lightweight models, integrating multi-modal data, and improving model interpretability.

Comparative Table

Year	Method	Architecture	Key Feature	Performance
2020	U-Net	CNN	Basic segmentation	Moderate
2020	3D U-Net	CNN	Multi-scale learning	High
2022	Attention U-Net	CNN + Attention	Boundary refinement	Improved
2022	Cascade U-Net	Multi-stage	ROI-based segmentation	High
2023	EfficientNet U-Net	Hybrid	Efficient scaling	Very High
2023	Hybrid Model	Seg + Classifier	End-to-end system	Excellent

Comparative Analysis
The comparative evaluation of renal tumor segmentation and classification techniques developed between 2020 and 2023 highlights a clear progression in deep learning methodologies, architectural sophistication, and performance outcomes. This evolution reflects the increasing complexity of medical imaging challenges and the growing need for accurate, reliable, and clinically deployable systems.

Evolution of Architectures: From Conventional CNNs to Hybrid Models
Early-stage approaches (2020) were primarily dominated by conventional convolutional neural network (CNN) architectures, particularly U-Net and 3D U-Net variants. These models established a strong foundation for biomedical image segmentation by leveraging encoder–decoder structures and skip connections to preserve spatial information. While these models achieved high Dice scores for kidney segmentation (often above 0.95), their performance in tumor segmentation was comparatively lower due to challenges such as irregular tumor boundaries, low contrast, and class imbalance.
The introduction of multi-scale learning significantly improved performance by enabling models to capture both fine-grained and coarse features. Multi-scale supervised 3D U-Net architectures demonstrated improved tumor segmentation accuracy, as they incorporated hierarchical feature representations and deep supervision. However, these improvements came at the cost of increased computational complexity and longer training times.
As research progressed into 2021–2022, there was a transition toward hybrid and enhanced architectures, incorporating attention mechanisms, cascaded frameworks, and boundary-aware designs. Attention-based models, such as Attention U-Net, enabled the network to focus on relevant tumor regions while suppressing irrelevant background information. This resulted in improved boundary delineation and reduced false positives.
Cascade architectures further enhanced segmentation by decomposing the problem into multiple stages. Typically, these models first segmented the kidney region and subsequently focused on tumor segmentation within the identified region of interest (ROI). This hierarchical approach effectively addressed class imbalance by narrowing the search space, leading to improved segmentation accuracy and computational efficiency.

Impact of EfficientNet Integration
One of the most significant advancements in recent years is the integration of EfficientNet as an encoder backbone within U-Net architectures. EfficientNet introduced a compound scaling method that optimizes network depth, width, and resolution simultaneously. This innovation enabled models to achieve higher accuracy with fewer parameters compared to traditional CNNs. EfficientNet-based U-Net models demonstrated superior feature extraction capabilities, particularly in capturing complex tumor textures and heterogeneous patterns. Comparative studies show that these models consistently outperform standard U-Net and V-Net architectures, achieving Dice scores above 0.90 and IoU values approaching 0.98 on benchmark datasets such as KiTS19 and KiTS21.

Another advantage of EfficientNet-based architectures is their parameter efficiency, which makes them more suitable for deployment in real-world clinical environments. Unlike heavy 3D CNN models, EfficientNet-based U-Net models strike a balance between accuracy and computational cost, enabling faster inference and reduced memory usage.

Role of Attention Mechanisms and Transformers

Attention mechanisms have played a crucial role in improving segmentation accuracy by enabling models to selectively focus on relevant regions. In renal tumor segmentation, attention modules help highlight tumor boundaries and suppress background noise, leading to more precise segmentation results.

Transformer-based architectures represent another major advancement in this domain. Unlike CNNs, which rely on local receptive fields, transformers use self-attention mechanisms to capture global contextual information. This capability is particularly useful for analyzing complex tumor structures that span across multiple regions of an image.

Models such as UNETR and Swin-UNet combine the strengths of CNNs and transformers, resulting in hybrid architectures that achieve state-of-the-art performance. However, these models require significant computational resources and large training datasets, which limit their applicability in resource-constrained settings.

Segmentation vs Classification: Toward Unified Frameworks

Another important trend observed in the literature is the shift from standalone segmentation models to integrated segmentation–classification frameworks. Early approaches treated segmentation and classification as separate tasks, requiring multiple models and processing steps. However, recent research focuses on developing end-to-end systems that perform both tasks simultaneously.

These unified frameworks improve diagnostic accuracy by leveraging shared feature representations. For instance, segmentation outputs can be used to guide classification models, enabling more accurate identification of tumor subtypes (benign vs malignant). Studies report classification accuracies exceeding 97% when segmentation and classification are integrated into a single pipeline.

This integration also reduces computational overhead and simplifies the workflow, making it more suitable for clinical deployment.

Uncertainty Estimation and Epistemic Neural Networks

A critical limitation of traditional deep learning models is their inability to quantify uncertainty in predictions. In medical applications, where decisions can have life-threatening consequences, understanding the reliability of model outputs is essential.

Epistemic neural networks address this challenge by modeling uncertainty arising from limited data and model parameters. Techniques such as Bayesian neural networks and Monte Carlo dropout enable the estimation of confidence intervals for predictions.

The inclusion of uncertainty estimation provides several advantages:

- **Improved reliability**: Clinicians can assess the confidence level of predictions
- **Error detection**: High uncertainty values can indicate potential misclassifications
- **Decision support**: Models can flag uncertain cases for further review

Recent studies highlight the importance of incorporating uncertainty estimation into medical imaging systems, particularly for safety-critical applications such as cancer diagnosis.

Performance Metrics and Benchmarking Trends

Performance evaluation in renal tumor segmentation is typically based on metrics such as:

- Dice Similarity Coefficient (DSC)
- Intersection-over-Union (IoU)
- Sensitivity and Specificity
- Hausdorff Distance (HD)

Comparative analysis shows that EfficientNet-based U-Net models consistently achieve higher Dice and IoU scores compared to traditional architectures. However, variations in dataset preprocessing, training strategies, and evaluation protocols make direct comparisons challenging.

Datasets such as KiTS19 and KiTS21 have become standard benchmarks, but their limited diversity raises concerns about model generalization. Models trained on these datasets may not perform well on real-world data with different imaging conditions.

Key Challenges Identified

Despite significant advancements, several challenges remain:

1. Class Imbalance

Tumor regions are significantly smaller than kidney regions, leading to biased predictions. Multi-stage and attention-based models partially address this issue but do not eliminate it entirely.

2. Data Scarcity and Variability

Limited availability of annotated medical datasets restricts model training. Variability in imaging protocols further complicates generalization.

3. Computational Complexity

Advanced models, particularly 3D CNNs and transformers, require high computational resources, limiting their practical deployment.

4. Lack of Standardization

Differences in evaluation metrics and experimental setups hinder fair comparison across studies.

5. Clinical Interpretability

Many deep learning models function as "black boxes," making it difficult for clinicians to trust their predictions.

Future Research Directions

Based on the comparative analysis, several promising research directions emerge:

- Development of lightweight and efficient architectures for real-time applications
- Integration of multi-modal imaging data (CT + MRI)
- Improved uncertainty estimation techniques
- Standardized benchmarking protocols
- Focus on explainable AI (XAI) for better clinical interpretability

Summary of Findings

The comparative analysis clearly demonstrates a transition from traditional CNN-based models to advanced hybrid architectures incorporating EfficientNet, attention mechanisms, and transformers. EfficientNet-based U-Net models currently represent the state-of-the-art due to their superior performance and efficiency.

Furthermore, the integration of epistemic neural networks has introduced a new dimension of reliability, making these systems more suitable for clinical applications. However, addressing challenges such as data scarcity, computational complexity, and generalization remains essential for the widespread adoption of these technologies.

Discussion

The development of deep learning-based renal tumor segmentation systems has significantly advanced over the past few years. EfficientNet-based U-Net architectures have become the preferred choice due to their balance between accuracy and computational efficiency. These models leverage compound scaling to improve feature extraction while maintaining manageable model complexity.

Attention mechanisms and transformer-based models have further enhanced segmentation performance by capturing both local and global features. Multi-stage architectures have also proven effective in improving tumor localization and addressing class imbalance.

The integration of epistemic neural networks represents a critical advancement in ensuring model reliability. By quantifying uncertainty, these models provide confidence measures that are essential for clinical decision-making. This feature is particularly important in medical imaging, where incorrect predictions can have serious consequences.

Despite these advancements, several challenges persist. Data scarcity and variability across datasets limit model generalization. Additionally, high computational requirements hinder real-time deployment in clinical environments. Future research should focus on developing lightweight models and integrating multi-modal data to improve performance and applicability.

Conclusion

This survey highlights the significant progress in renal tumor segmentation and classification using deep learning techniques. EfficientNet-based U-Net architectures have emerged as state-of-the-art models, providing superior performance in feature extraction and segmentation accuracy. The integration of attention mechanisms, multi-stage frameworks, and hybrid architectures has further enhanced model capabilities.

Epistemic neural networks add a critical dimension by enabling uncertainty estimation, improving the reliability and trustworthiness of predictions. These advancements make deep learning models increasingly suitable for clinical applications.

However, challenges such as data imbalance, computational complexity, and limited generalization remain. Addressing these challenges requires the development of efficient architectures, improved training strategies, and standardized evaluation frameworks.

Future research should focus on creating interpretable, scalable, and clinically deployable systems. The integration of segmentation, classification, and uncertainty estimation into unified frameworks holds great promise for advancing renal cancer diagnosis and treatment.

References

Zhao, W., et al. (2020). Multi-scale supervised 3D U-Net. https://doi.org/10.48550/arXiv.2004.08108

Qin, T., et al. (2020). Reinforcement learning augmentation. https://doi.org/10.48550/arXiv.2002.09703

Türk, F., et al. (2020). Hybrid V-Net. *Mathematics*, 8(10). https://doi.org/10.3390/math8101772

Hu, S., et al. (2022). Boundary-aware network. https://doi.org/10.48550/arXiv.2208.13338

Lin, C., et al. (2022). Cascade segmentation. https://doi.org/10.1007/978-3-031-16440-8_12

Abdelrahman, A., & Viriri, S. (2022). Kidney tumor segmentation survey. https://doi.org/10.1109/ACCESS.2022.3145678

Isensee, F., et al. (2021). nnU-Net. *Nature Methods*. https://doi.org/10.1038/s41592-020-01008-z

Tan, M., & Le, Q. (2019). EfficientNet. https://doi.org/10.48550/arXiv.1905.11946

Ronneberger, O., et al. (2015). U-Net. https://doi.org/10.1007/978-3-319-24574-4_28

Oktay, O., et al. (2018). Attention U-Net. https://doi.org/10.48550/arXiv.1804.03999

Milletari, F., et al. (2016). V-Net. https://doi.org/10.1109/3DV.2016.79

Zhou, Z., et al. (2019). UNet++. https://doi.org/10.1109/TMI.2018.2889096

Chen, J., et al. (2021). Transformer segmentation. https://doi.org/10.1016/j.media.2021.102136

Dosovitskiy, A., et al. (2021). Vision Transformer. https://doi.org/10.48550/arXiv.2010.11929

Hatamizadeh, A., et al. (2022). UNETR. https://doi.org/10.48550/arXiv.2103.10504

Cao, H., et al. (2022). Swin-Unet. https://doi.org/10.48550/arXiv.2105.05537

Myronenko, A. (2019). 3D MRI segmentation. https://doi.org/10.48550/arXiv.1810.11654

Sudre, C., et al. (2017). Generalised Dice loss. https://doi.org/10.48550/arXiv.1707.03237

Kendall, A., & Gal, Y. (2017). Uncertainty in deep learning. https://doi.org/10.48550/arXiv.1703.04977

Gal, Y., & Ghahramani, Z. (2016). Bayesian dropout. https://doi.org/10.48550/arXiv.1506.02142