# Medical Image Processing: Automated Analysis of Depiction of Renal Cancers in CT Images

**Student**: Zixin Hao (1309180)

**Institution**: Faculty of Engineering and IT

**Supervisor**: Brian Chapman

**Date**: 11/03/2024

# Table of Contents

# Abstract

This research focuses on leveraging deep learning, particularly convolutional neural networks (CNN), to enhance the automatic segmentation of kidney tumors in CT images. Early detection and treatment of kidney cancer are crucial, and this study aims to improve the quantification and treatment decision-making through semantic segmentation. Despite deep learning's potential in medical image analysis, challenges like insufficient data, weak generalization, and high computational demands persist. The project will develop a fully automatic, efficient, and lightweight model with associated strategies to boost segmentation accuracy, robustness, and efficiency. Utilizing nnU-net as the core, it will explore integrating EfficientNet or EfficientUnet++ for better model performance and address issues like foreground/background imbalance and enhancing recall rates. The goal is to devise new architectures and strategies to overcome existing hurdles, aiming for applications in early cancer diagnosis and treatment. The study will employ public datasets and various computing resources. The timeline covers literature review, model experimentation, and final report preparation.

# 1. Introduction

Many factors have led to a gradual increase in the incidence of cancer among humans, necessitating early detection and treatment for effective disease prevention. Kidney cancer ranks as the ninth most common cancer among men and the fourteenth among women. The chances of a cure are significantly higher if kidney cancer is detected and treated in its early stages [1]. Often, kidney cancer does not show symptoms initially and is typically found incidentally; the diagnosis is commonly made through a biopsy and the surgical removal of the tumour. Tumour removal can be performed as a radical nephrectomy or as a partial nephrectomy. Due to its numerous advantages, the latter has become the preferred approach.

To enhance the reliability in quantifying tumour characteristics and facilitate accurate comparisons in treatment decisions for kidney tumors, various nephelometry scoring systems have been introduced. Semantic segmentation offers a detailed quantitative analysis of the tumour and the organ involved. Therefore, it plays a critical role in a wide range of clinical applications [2].

However, automatic segmentation remains a challenge due to the heterogeneity of renal tumors and low contrast with the surrounding tissue. In recent years, deep learning methods, particularly CNN, have shown great potential in the field of medical image analysis [3]. These methods are able to learn complex features from large amounts of data, improving segmentation accuracy and efficiency. However, existing research still faces problems such as insufficient data, weak algorithm generalization capabilities, and high computing resource requirements. In addition, most studies overfit specific tasks and data sets, focusing too much on scores and ignoring efficiency, which violates the original intention of clinical applications.

**The scope of this research is to improve renal tumor segmentation in CT images by developing a fully automatic, efficient and lightweight model, leveraging a set of related strategies (such as pre/post-processing). In order to move closer to clinical application, not only accuracy, but also generalization ability and efficiency are considered.**

# 2. Literature Review

## 2.1 CNN

CNNs are a significant branch in the field of deep learning, having achieved tremendous success in image processing and recognition tasks since their introduction by LeCun et al. in 1998 [3]. The core advantage of CNNs lies in their ability to automatically learn image features through convolutional layers, making them exceptionally effective in processing data with spatial hierarchies. CNNs, through the stacking of multiple convolutional and pooling layers, gradually abstract more complex image content from simple edges and texture features, which is particularly important in medical image analysis. With the development of CNNs, many critical derivative networks have emerged. For example, VGGNet improved performance by increasing network depth [4]; ResNet introduced residual connections to solve the degradation problem in deep network training [5];

## 2.2 U-net

In medical image segmentation, the most commonly used and popular architectures are U-net and various variant networks (Particularly Cascaded U-net, nnU-net, 3D U-net).

The introduction of the U-net architecture marked a significant milestone in medical image analysis, especially in the field of image segmentation. Proposed by Ronneberger et al. in 2015, U-net is a network structure specifically designed for medical image segmentation [6]. Characterized by its symmetric U-shaped structure, U-net effectively combines deep and shallow features in image segmentation. U-net uses skip connections to merge features from different levels, maintaining high accuracy in details, which is crucial for processing complex medical images.

The U-Net architecture, introduced by Ronneberger and colleagues in 2015, along with its 3D adaptations by Milletari et al. and Çiçek et al., stand as foundational techniques for deep learning-driven segmentation of medical imagery [6, 7]. Following their introduction, a myriad of enhancements to U-Net have been developed, often yielding significant performance boosts compared to the original U-Net framework. In the KiTS challenge at MICCAI2019, the top 15 techniques were derivatives of the (3D) U-Net architecture from 2016, underscoring its significance in the realm of biomedical image segmentation [8, 9].

In 2018, Isensee and associates achieved leading-edge results in multiple prestigious 3D segmentation competitions by employing a basic U-Net in conjunction with a novel approach for optimizing a limited range of hyperparameters and preprocessing techniques [10, 11]. Most of the top-ranked teams in the KiTS19/21/23 challenge use this network and cascaded network as the cornerstone.

# 3. Expected Results

**Expected Results:** New architecture and a set of related strategies

The application of **lightweight** and **efficient** deep learning models in medical image analysis is of great significance, especially in **reducing reliance on data annotation and enhancing model robustness and efficient**. **My research goal is to design new architecture and strategies (e.g. data preparation and processing, tuning strategies beyond the model) to address these problems, resulting in medical image segmentation** pipelines that are less reliant on large, labeled data sets for training and are more robust to changes in data sets. This pipeline could form the foundations for future work, beyond the scope of this proposal, such as an advanced end-to-end computer-aided diagnosis (CAD) system (Integrate kidney tumor segmentation and classification). There is a correlation between the two visual tasks of tumor area segmentation and classification. Tumor segmentation is closely related to features such as tumor morphology. Therefore, tumor segmentation can guide feature extraction for tumor classification, thereby improving performance.

# 4. Methodology

This project will use an unoptimized network (such as U-net, nnU-net, 3D U-net, and EfficientUNet++) as the baseline. For the new network, I will choose nnU-net as the backbone network because of its demonstrated effectiveness in the KiTS19 competition: the leading fifteen methods from the (3D) U-Net model introduced in 2016, highlighting its foundational role in shaping the domain of biomedical image segmentation [11]. (2) nnU-Net performed best in the KiTs19 Challenge [12, 13].

I will seek improvements to the nnU-net architecture by exploring the following potential modifications:

- **Improving efficiency**: Integrate EfficientNet or EfficientUnet++ into nnU-net to try to increase efficiency. (Abdelrahman A. and Viriri S. point out EfficientNets' efficiency in kidney cancer and CT scan classification, highlighting their simplicity, quicker training, and stable accuracy in their paper [14]. The combination of these two excellent networks holds great potential)
- **Resolving foreground/background imbalance issues [14]:** Considering cascaded segmentation
- **Improving recall rate and surface dice:** The competing teams' methods demonstrated higher precision than recall when segmenting tumors, indicating that it is often difficult to find the entire tumor. Surface dice is generally low, so we should find ways to strengthen edge segmentation [8].
- **Improving the receptive field of nnU-net and using complete images to train the 3D U-net in nnU-net:** Isensee F et al. mentioned in the paper that the field of view of nnU-net is limited, and due to the impact of GPU memory, the 3D U-net in nnU-net can only be trained on image patches [10].
- **Designing out-of-model strategies:** Many papers and experiments show that the original U-net and nnU-net are more robust and perform better than the variants [8, 10,12]. What is most important is not the design of the network [15]. Isensee F et al. also believe that many studies overfit to a single dataset because the quality of the hyperparameter configuration often masks the effect of the evaluated architectural modifications [11]. There are no common architectural modifications among the top 15 rankings in the KiTS19 challenge [8, 11]. Therefore, I believe the non-model aspects of segmentation methods are more influential and are also severely underestimated.

**Ultimately, to evaluate the robustness of the proposed method, it will be subjected to testing across multiple datasets and thorough ablation studies will be conducted as part of this research proposal.**

## 4.1 Data

The research will use public data sets, such as KiTS23, TCGA-KIRC [16, 17]. Data sampling encompasses sources from over 50 hospitals, indicating a notable level of diversity [8]. Both CT image collection and Annotation are in line with the scientific approach. However, it is important to note that in KiTS23, there is a bias in the data because the patients sampled represent only a subset of all patients seen for concern of renal malignancy [8]. Additionally, the exclusion of cysts contributed to (but did not fully account for) the lower proportion of lesions postoperatively found to be benign in our cohort compared to what has been reported elsewhere (8% vs 30%) [8].

## 4.2 Evaluation

- **Performance metrics:** The proposal plans to adopt Dice Score and Surface Dice Score as the principal evaluation metrics to precisely gauge the model's performance. The combination of these two evaluation indicators can not only comprehensively evaluate the overall segmentation performance of the model, but also accurately reflect the model's ability to process details.

- **Train/test paradigm:** To guarantee precise assessment of model performance and mitigate overfitting issues, cross-validation (CV) is utilized as the fundamental strategy for model evaluation.

## 4.3 Computing resources

Three candidate ways:
1. GPU Cloud Computing platform (Rent a RTX4090 GPU instance)
2. Local GeForce RTX 3060 Laptop GPU (If the EfficientNet family U-net doesn't require much accelerating)
3. Google TPU platform

# 5. Timetable

The following Table 1. shows the research plan.

**Table 1**. *Research Plan*

| Stages | Work content (high level) |
|---|---|
| 1. Week 1 - Week2 (10/03/2024) | Literature Review |
| 2. Week 3 (15/03/2024) | Finished Formal Research Proposal |
| 3. Week4 - Week5 (31/03/2024) | Learn and try models(baseline): EfficientUNet, nnUnet, EfficientUNet++, "An attempt at beating the 3D U-Net" |
| 4. Mid-break-Week6(14/04/2024) | Try improvement point: 1. Architecture modifications (add modules); 2. I want to combine EfficientNet and nnUnet. 3. Look for improvements outside of the model (multiple papers confirm its importance) |
| 5. Week 7(21/04/2024) | Buffer week; finishing unfinished experiment |
| 6. Week8 - week 9 (5/05/2024) | Start writing; Do supplementary experiments (for charts required in report) |
| 7. Week10-Week11(19/05/2024) | Finished final report |
| 8. Week 12 (26/05/2024) | Prepare Oral presentations |

# 6. Thesis Outline or Structure

The Figure 1. presents the thesis outline at a conceptual level, offering an abstract overview of the framework.

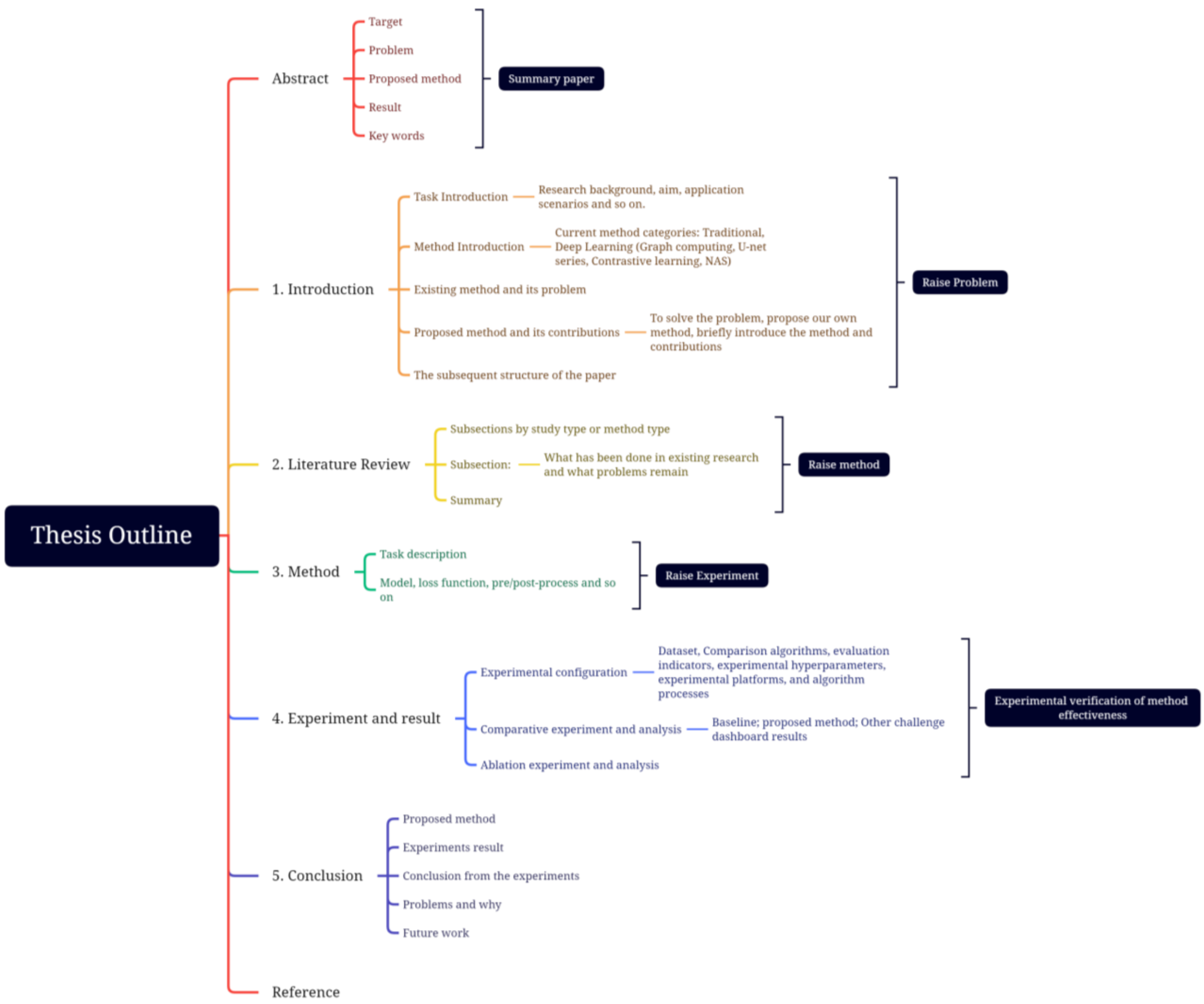


Figure 1. Thesis Outline

# Reference List


[1] Hsiao C H, Lin P C, Chung L A, et al. A deep learning-based precision and automatic kidney segmentation system using efficient feature pyramid networks in computed tomography images[J]. Computer Methods and Programs in Biomedicine, 2022, 221: 106854.
[2] Carmi S, Hui K Y, Kochav E, et al. Sequencing an Ashkenazi reference panel supports population-targeted personal genomics and illuminates Jewish and European origins[J]. Nature communications, 2014, 5(1): 4835.
[3] LeCun Y. The MNIST database of handwritten digits[J]. http://yann. lecun. com/exdb/mnist/, 1998.
[4] Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition[J]. arXiv preprint arXiv:1409.1556, 2014.
[5] He K, Zhang X, Ren S, et al. Deep residual learning for image recognition[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2016: 770-778.
[6] Ronneberger O, Fischer P, Brox T. U-net: Convolutional networks for biomedical image segmentation[C]//Medical image computing and computer-assisted intervention–MICCAI 2015: 18th international conference, Munich, Germany, October 5-9, 2015, proceedings, part III 18. Springer International Publishing, 2015: 234-241.
[7] Milletari F, Navab N, Ahmadi S A. V-net: Fully convolutional neural networks for volumetric medical image segmentation[C]//2016 fourth international conference on 3D vision (3DV). Ieee, 2016: 565-571.
[8] Heller N, Isensee F, Maier-Hein K H, et al. The state of the art in kidney and kidney tumor segmentation in contrast-enhanced CT imaging: Results of the KiTS19 challenge[J]. Medical image analysis, 2021, 67: 101821.
[9] Isensee F, Jäger P F, Kohl S A A, et al. Automated design of deep learning methods for biomedical image segmentation[J]. arXiv preprint arXiv:1904.08128, 2019.

[10] Isensee F, Petersen J, Klein A, et al. nnu-net: Self-adapting framework for u-net-based medical image segmentation[J]. arXiv preprint arXiv:1809.10486, 2018.
[11] Isensee F, Jaeger P F, Kohl S A A, et al. nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation[J]. Nature methods, 2021, 18(2): 203-211.
[12] Liu J, Yildirim O, Akin O, et al. AI-driven robust kidney and renal mass segmentation and classification on 3D CT images[J]. Bioengineering, 2023, 10(1): 116.
[13] Heller N, Isensee F, Trofimova D, et al. The KiTS21 Challenge: Automatic segmentation of kidneys, renal tumors, and renal cysts in corticomedullary-phase CT[J]. arXiv preprint arXiv:2307.01984, 2023.
[14] Abdelrahman A, Viriri S. EfficientNet Family U-Net Models for Deep Learning Semantic Segmentation of Kidney Tumors on CT Images[J]. Frontiers in Computer Science, 2023, 5: 1235622.
[15] Litjens G, Kooi T, Bejnordi B E, et al. A survey on deep learning in medical image analysis[J]. Medical image analysis, 2017, 42: 60-88.
[16] Heller N, Sathianathen N, Kalapara A, et al. The kits19 challenge data: 300 kidney tumor cases with clinical context, ct semantic segmentations, and surgical outcomes[J]. arXiv preprint arXiv:1904.00445, 2019.
[17] Akin, O., Elnajjar, P., Heller, M., Jarosz, R., Erickson, B. J., Kirk, S., Lee, Y., Linehan, M. W., Gautam, R., Vikram, R., Garcia, K. M., Roche, C., Bonaccio, E., & Filippini, J. (2016). The Cancer Genome Atlas Kidney Renal Clear Cell Carcinoma Collection (TCGA-KIRC) (Version 3) [Data set]. The Cancer Imaging Archive. https://doi.org/10.7937/K9/TCIA.2016.V6PBVTDR