

“Segmenting Kidney Tumors” Rubric

DS 4002 – Spring 2025 - Instructor: Ali Nilforoush

Due: TBD

Submission format: Upload a link to your repository on Canvas

Individual Assignment

General Description: Submit to Canvas a link to your completed case study repository.

Preparatory Assignments – Everything you have learned thus far as a 2nd year data science student.

Why am I doing this? This is a way to apply your data science skill set to improving healthcare and an opportunity to work with and implement convolutional neural networks (CNNs).

What am I going to do? You will follow the dataset import and processing instructions in the GitHub repository found here (<https://github.com/ali-rn/DS4002-CS3/>). You will build a CNN model and train it on the dataset to segment the kidneys, kidney tumor, cyst, and background. You will make a figure depicting your architecture and provide statistical t-Test comparisons to two other premade models based on three metrics: dice similarity score (DSC), precision, recall. You will provide a .pdf report with your architecture visualization, visualizations of your results, and written descriptions of each component. The written descriptions are meant to demonstrate your understanding of how your model works and what the metrics mean. You will also provide your model, training, and testing code in a zip file.

Tips for success:

- Use the example model code and follow the reproduction steps to better understand the workflow. This example model is what you will be comparing your own model to.
- Reach out to the teaching team of your class for any help along the way. Be sure to ask for help early and often!
- Use Weights & Biases’ hyperparameter sweep to optimally train your model

How will I know I have Succeeded? You will meet expectations on this case study when you follow the criteria in the rubric below.

Formatting	<ul style="list-style-type: none">• Summary<ul style="list-style-type: none">○ Architecture (Method)○ Results○ Discussion• Code

Summary (*.pdf)	<p><u>Goal:</u> Write a summary about your CNN model and results</p> <ul style="list-style-type: none"> • Methods: <ul style="list-style-type: none"> o Provide a visualization of your model architecture, including the encoder and decoder side o Describe how each part works and the role the layers play in your overall model o Justify any decisions you made and your model design approach o Include train-test split, hyperparameters (i.e., num epochs, learning rate, optimization, loss function) • Results: <ul style="list-style-type: none"> o Provide visualizations of your comparison with the two base models along DSC, precision, recall, and any other metrics you deem fit o Explain and interpret each of the metrics • Discussion <ul style="list-style-type: none"> o Explain your results and how the models compare o What changes could you make to improve your model?
Code (*.zip)	<p><u>Goal:</u> Document the code of your CNN model testing and training</p> <ul style="list-style-type: none"> • Use the two provided models to inform your training and testing workflow • Tweak the starter scripts and data processing scripts as you see fit (hint: you may need to incorporate more data to train your models properly!) • Comment your code well! • All scripts should be zipped into a single folder to be submitted alongside your summary

Acknowledgements: Special thanks to Jess Taggart from UVA CTE for coaching on making this rubric. This structure is pulled from [Streifer & Palmer \(2020\)](#).