DEPARTMENT OF INFORMATICS

TECHNISCHE UNIVERSITÄT MÜNCHEN

Clinical Project

Prediction of MRI-based Lesion Masks from CT Perfusion Data

Felician Schwarz

DEPARTMENT OF INFORMATICS

TECHNISCHE UNIVERSITÄT MÜNCHEN

Clinical Project

Prediction of MRI-based Lesion Masks from CT Perfusion Data

Author:	Felician Schwarz
Supervisor:	Mina Nessiem
Advisor:	Advisor
Submission Date:	Submission date

Acknowledgments

Abstract

Contents

1 Introduction

Stroke is one of the leading causes of death and disability worldwide, with ischemic stroke accounting for approximately 87% of all cases. In the context of stroke treatment, the phrase "Time is Brain" encapsulates a fundamental principle: time is a crucial factor that significantly influences the extent of irreversible brain damage and the potential for recovery. Each minute delay in treatment results in the loss of approximately 1.9 million neurons, emphasizing the critical nature of rapid intervention.

The challenge in treating ischemic stroke lies in accurately differentiating between irreversibly damaged tissue (the infarct core) and potentially salvageable tissue (the penumbra). As time progresses, the core expands at the expense of the penumbra, highlighting the urgency of accurate diagnostic imaging and prompt treatment decisions.

Currently, computed tomography (CT) is the primary imaging modality used in the acute phase of stroke due to its widespread availability, speed of acquisition, and ability to differentiate between ischemic and hemorrhagic stroke. Perfusion CT, which measures blood flow in the brain over time using contrast agents, plays a crucial role in identifying the core and penumbra regions. However, magnetic resonance imaging (MRI) provides superior soft tissue contrast and more accurate visualization of the final infarct volume, making it the preferred method for follow-up assessment.

This clinical gap—between the initial CT imaging used for urgent decision-making and the subsequent MRI that better represents the actual tissue outcome—presents both a challenge and an opportunity for computational approaches. The ISLES 2024 dataset, the first longitudinal multimodal multi-center real-world dataset in (sub-)acute stroke, offers a unique resource to address this gap by providing paired acute CT and follow-up MRI data from the same patients.

1.1 Project Objectives

The primary goal of this clinical project is to predict MRI-based lesion masks from CT data, effectively bridging the temporal and informational gap between acute CT imaging and follow-up MRI. Specifically, this project aims to:

- Analyze and preprocess imaging data from the ISLES 2024 dataset, including CT perfusion maps (CBF, CBV, MTT, TTP) and associated clinical information.

- Develop and evaluate computational models that can predict final infarct volumes (as represented in MRI-derived lesion masks) based on sub-acute CT data.
- Assess the generalizability of the developed models across different clinical centers and varying patient characteristics.
- Explore the potential clinical utility of such predictive models in acute stroke management decision-making.

The project builds upon recent advances in deep learning approaches to medical image segmentation, particularly in the context of stroke imaging, while addressing the specific challenges posed by cross-modality prediction and variability in imaging parameters across clinical centers.

2 Theoretical Foundations

This chapter provides the theoretical background necessary to understand the clinical and technical aspects of stroke imaging and analysis.

2.1 Ischemic Stroke Pathophysiology

In ischemic stroke, a blood clot obstructs blood flow to a region of the brain, leading to oxygen and nutrient deprivation in the affected area. The brain tissue affected by a stroke can be categorized into two main regions:

- **Infarct Core**: The central region of severely reduced blood flow where brain cells have already died or are irreversibly damaged.
- **Penumbra**: The surrounding area of compromised but still viable brain tissue, which can potentially be saved with timely intervention.

The concept "Time is Brain" emphasizes that as time progresses, the infarct core expands at the expense of the penumbra, making rapid diagnosis and treatment essential to preserve brain function.

2.2 Imaging Modalities in Stroke

2.2.1 Computed Tomography (CT)

CT imaging is the first-line imaging modality in acute stroke care due to several advantages:

- Faster acquisition (minutes vs. 30-45 minutes for MRI)
- Widespread availability in hospitals
- No contraindications related to metal implants or pacemakers
- No requirement for patients to remain completely still

- Cost-effectiveness
- Ability to quickly differentiate between ischemic and hemorrhagic stroke

However, CT provides primarily structural and anatomical information and is less sensitive than MRI for detecting early ischemic changes.

2.2.2 Magnetic Resonance Imaging (MRI)

MRI is commonly used for follow-up imaging due to:

- Superior soft tissue contrast
- More accurate visualization of the final infarct volume
- Ability to detect even small ischemic lesions

Diffusion-weighted imaging (DWI) is particularly valuable, as it is the only imaging technique that reliably demonstrates parenchymal injury within minutes to hours from stroke onset.

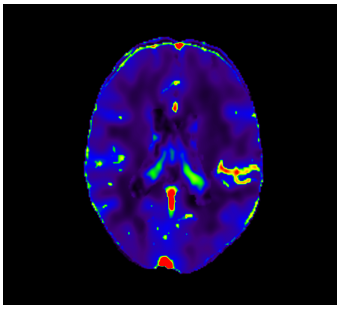
2.3 Perfusion CT Analysis

2.3.1 Perfusion Maps

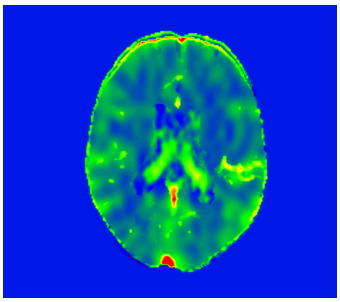
Perfusion maps and imaging techniques:

- Derived volumetric data visualizing blood perfusion parameters, not raw CT.
- CBF (Cerebral Blood Flow): Blood flow per time through a brain volume (ml/100g tissue/min).
- CBV (Cerebral Blood Volume): Total blood volume in a brain region (ml/100g tissue).
- MTT (Mean Transit Time): Average time for blood to pass through a region.
- TTP (Time to Peak): Time to maximum contrast agent concentration in an area.
- Tmax (Time to Maximum): Delay until peak of tissue residue function, sensitive indicator for hypoperfusion.
- CTA (Computed Tomography Angiography): CT imaging with contrast agents to visualize vascular structures and detect pathologies.

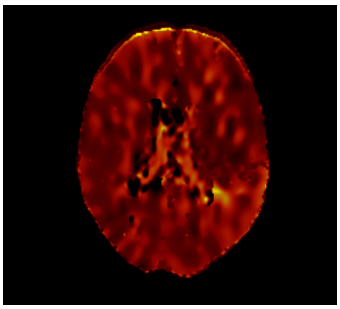
- CTP (Computed Tomography Perfusion): Dynamic CT method quantifying brain perfusion through temporal contrast distribution analysis.



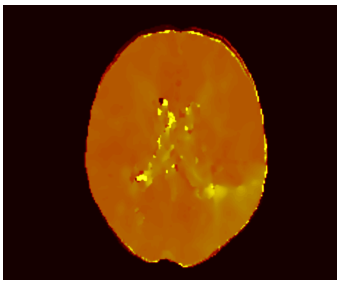
(a) Cerebral Blood Flow (CBF) map



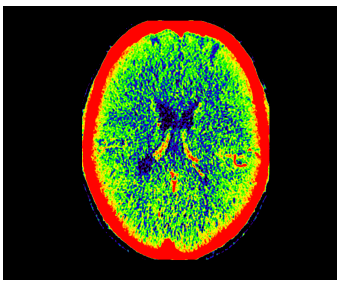
(b) Cerebral Blood Volume (CBV) map



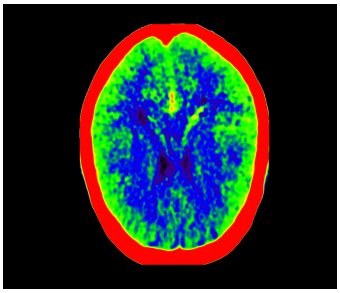
(c) Mean Transit Time (MTT) map



(d) Time to Maximum (Tmax) map



(e) CT Angiography (CTA) map



(f) CT Perfusion (CTP) map

Figure 2.1: Different perfusion maps used in ischemic stroke analysis

2.3.2 Technical Considerations

Perfusion CT measures blood flow in the brain over time (every 1-2 seconds over 45-60 seconds) using a contrast agent. The derived perfusion maps help differentiate between irreversibly damaged (core) and salvageable (penumbra) tissue.

Different software packages use different deconvolution algorithms and thresholds, which can lead to varying estimations of core and penumbra volumes. This variability presents a challenge for standardization across clinical centers and research studies.

2.4 Lesion Masks and Ground Truth

Lesion masks represent binary infarct masks in the dataset and are derived from MRI images. They show the actual final infarct areas and are co-registered to the Non-Contrast CT (NCCT). These masks serve as ground truth—the target that predictive models aim to accurately reproduce based on earlier CT data.

3 Data Preprocessing

3.1 Dataset

The dataset contains a file named isles24_multimodal_5fold_NIHSSstratified.json which contains some useful information about the Dataset. It is divided into five folds as shown by the "fold" field for each case. The split is stratified by NIHSS scores, meaning that each fold has a balanced distribution of stroke severity, which ensures that cross-validation results are robust and representative of the entire cohort.

3.1.1 Dataset Structure

The dataset is organized into three main subdirectories: "derivatives", "phenotype", and "raw_data".

- "raw_data" contains the original data from the ISLES24 challenge.
- "phenotype" contains clinical and demographic information about the patients.
- "derivatives" contains the processed data for each patient.

The tree subdirectories have a similar structure to the following example of derivatives for patient SUB-STROKE0001: Since the folder structure is somewhat complex, we use the `tree path /F` command to visualize the structure of one patient's data (stored in SUB-STROKE0001). Each patient in the dataset has a similar directory structure:

```
D:\TUM_CLINICAL_PROJECT\ISLES24_COMBINED\DERIVATIVES\SUB-STROKE0001
+---ses-01
|   |   sub-stroke0001_ses-01_space-ncct_cta.nii.gz
|   |   sub-stroke0001_ses-01_space-ncct_ctp.nii.gz
|   |
|   \---perfusion-maps
|           sub-stroke0001_ses-01_space-ncct_cbf.nii.gz
|           sub-stroke0001_ses-01_space-ncct_cbv.nii.gz
|           sub-stroke0001_ses-01_space-ncct_mtt.nii.gz
|           sub-stroke0001_ses-01_space-ncct_tmax.nii.gz
```

```
|
\---ses-02
            sub-stroke0001_ses-02_lesion-msk.nii.gz
```

As shown above, the patient data is organized into two sessions:

- **ses-01**: Contains the CT angiography (CTA) and CT perfusion (CTP) scans, acquired in the acute phase, along with derived perfusion maps (cerebral blood flow, cerebral blood volume, mean transit time, and time-to-maximum).
- **ses-02**: Contains the lesion mask derived from follow-up MRI, which serves as the ground truth for our prediction task.

This structure follows the BIDS (Brain Imaging Data Structure) format, which is a standardized way of organizing neuroimaging data. The naming convention includes subject ID, session, image space, and modality, making it easier to identify and process the files programmatically.

3.2 Registration

Even though the data has already been registirated we double check this by writing a short script. We compare the affine matrix of the first images of the CT, perfusion and lesion. We get the following results:

```
Filename: sub-stroke0001_ses-01_space-ncct_cta.nii.gz
(-120.0, 254.75277709960938, -58.179813385009766)
(0.46875, 0.46875, 2.0)
(1.0, 0.0, 0.0, 0.0, -0.9659258450112692, -0.25881896370865276, 0.0,
    0.25881897523030545, -0.9659258480984858)
(512, 595, 75)
Filename: sub-stroke0001_ses-01_space-ncct_ctp.nii.gz
(-120.0, 216.44757080078125, -201.1368408203125)
(0.46875, 0.46875, 2.0)
(1.0, 0.0, 0.0, 0.0, 1.0, 0.0, 0.0, 0.0, 1.0)
(512, 595, 75)
Filename: sub-stroke0001_ses-02_lesion-msk.nii.gz
(-120.0, 254.75277709960938, -58.17981719970703)
(0.46875, 0.46875, 2.0)
(1.0, 0.0, 0.0, 0.0, -0.9659258410376269, -0.2588189915146001, 0.0,
    0.25881899006014436, -0.9659258406479068)
(512, 595, 75)
Filename: sub-stroke0001_ses-01_space-ncct_cbf.nii.gz
(-120.0, 254.75277709960938, -58.17981719970703)
```

```
(0.46875, 0.46875, 2.0)
(1.0, 0.0, 0.0, 0.0, -0.9659258410376269, -0.2588189915146001, 0.0,
    0.25881899006014436, -0.9659258406479068)
(512, 595, 75)
Filename: sub-stroke0001_ses-01_space-ncct_cbv.nii.gz
(-120.0, 254.75277709960938, -58.17981719970703)
(0.46875, 0.46875, 2.0)
(1.0, 0.0, 0.0, 0.0, -0.9659258410376269, -0.2588189915146001, 0.0,
    0.25881899006014436, -0.9659258406479068)
(512, 595, 75)
Filename: sub-stroke0001_ses-01_space-ncct_mtt.nii.gz
(-120.0, 254.75277709960938, -58.17981719970703)
(0.46875, 0.46875, 2.0)
(512, 595, 75)
Filename: sub-stroke0001_ses-01_space-ncct_tmax.nii.gz
(-120.0, 254.75277709960938, -58.17981719970703)
(0.46875, 0.46875, 2.0)
(1.0, 0.0, 0.0, 0.0, -0.9659258410376269, -0.2588189915146001, 0.0,
    0.25881899006014436, -0.9659258406479068)
(512, 595, 75)
```

We see that the affine matrix is the same for all images except for the lesion mask. (This is fine as the lesion mask is not registered)

3.3 Normalization

Medical images from different sources can have significant variations in intensity values due to differences in imaging protocols, scanner hardware, and acquisition parameters. To address this variability and improve the performance of machine learning models, we apply Z-score normalization to our dataset.

3.3.1 Z-Score Normalization

Z-score normalization (also known as standardization) transforms the intensity values of an image to have a mean of 0 and a standard deviation of 1. For each voxel value x in the dataset, the normalization formula is:

$$z = \frac{x - \mu}{\sigma} \tag{3.1}$$

where μ is the mean intensity value of the entire volume and σ is the standard deviation.

This normalization technique offers several advantages for medical image analysis:

- It centers the data distribution around zero, making different scans more comparable
- It helps deep learning models converge faster during training
- It reduces the influence of variations in scanner settings and image acquisition parameters

3.3.2 Normalization Script

To make sure that we work with normalized data, I developed a Python script that iterates over all the NIfTI files in our dataset and applies Z-score normalization to them. In addition, the script visualizes the original and normalized images side by side and the difference map.

Figure 3.1 demonstrates the effects of Z-score normalization on a CT scan slice. This visualization highlights several important observations about the normalization process:

As clearly illustrated in Figure 3.1, Z-score normalization has several observable effects:

- **Range transformation:** The original image has intensity values ranging from approximately -1027 to 1585, while the normalized image values are scaled to approximately -0.63 to 4.01. This demonstrates how the normalization process compresses the wide range of CT Hounsfield units into a standardized range centered around zero.
- **Statistical properties:** The mean value of the original image is -592.78 with a standard deviation of 590.12. After normalization, the mean is much closer to zero (0.14) with a standard deviation close to one (1.05), which closely aligns with the goal of Z-score normalization.
- **Contrast preservation:** Despite the intensity value transformation, the normalized image preserves the important structural features and contrast relationships of the original image.
- **Difference map:** The difference image (top right) highlights where the most significant changes occurred. The skull region shows the most pronounced transformation (in red), which is expected given that bone has the highest Hounsfield unit values in CT scans.
- **Histogram redistribution:** The bottom row histograms reveal how normalization transforms the intensity distribution. While maintaining the general bimodal

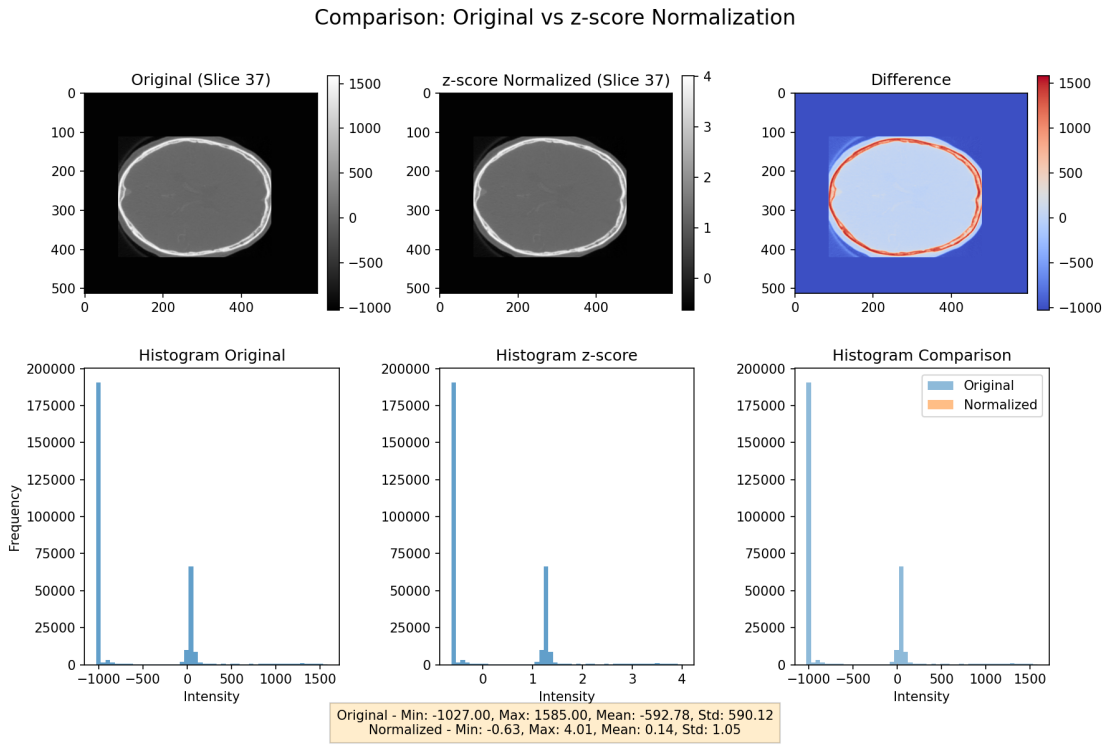


Figure 3.1: Example of Z-score normalization applied to a CT image (slice 37). Top row: Original image (left), normalized image (center), and difference map (right). Bottom row: Corresponding histograms showing intensity distributions before and after normalization. Note the significant change in intensity range and scale after normalization.

shape (representing primarily air/background and soft tissue/bone), the distribution is rescaled to center around zero, making it more suitable for deep learning algorithms.

This standardization process ensures that intensity values across different scans become comparable, which is particularly important when training machine learning models on datasets acquired from different scanners or institutions.

Note: I created this script to work with my local directory structure. If you wish to use this script, please modify the file paths to match your directory structure.

4 UNet Test Model Implementation with MONAI

To gain a better understanding of the data and task at hand, we first implement a simple UNet model using the MONAI framework. This serves as a baseline implementation before developing our own custom model architecture in later stages.

4.1 MONAI Pipeline Overview

4.2 UNet Architecture

4.3 Implementation

The Python libraries MONAI and PyTorch offer a wealth of features that facilitate the implementation of a UNet model. One useful feature is `monai.transforms`, which provides a range of transformations for data preprocessing.

To train a UNet model on volumetric data using MONAI, it is essential that all inputs in a batch share the same spatial dimensions. If the spatial dimensions differ, stacking the data into a 5D tensor will fail. Additionally, the image and its corresponding mask must be congruent in shape and distance. Therefore, a preprocessing pipeline is required to:

1. Unify spacings by resampling to a uniform voxel spacing
2. Center and crop the data to a common region of interest (ROI)
3. Apply padding to ensure a fixed size

Furthermore, it is necessary to apply these transformations synchronously to both the image and the mask.

For example:

- Voxel spacing is unified using the transformation `Spacingd(keys=["image", "label"], pixdim=[1.0, 1.0, 1.0])`, ensuring that all voxels have the same

spacing along each axis (the physical distance between the centers of adjacent voxels)

- The voxel shape must be divisible by a factor k. This is enforced by: `DivisiblePadd(keys=["image", "label"], k=16)`

We begin by implementing a minimal viable model that uses the Binary Cross-Entropy (BCE) loss function and the Dice metric. The model is trained on a single batch for five epochs, taking approximately 10 minutes. The results after training are:

- Loss: 0.1655
- Dice: 0.0000

A Dice score of 0 implies that the model is not matching the target masks at all, which likely means the model is either not detecting any relevant features or making entirely incorrect predictions. To further investigate, we visualize both the predicted mask and the ground truth mask. As shown in Figure 4.1, the predicted mask does not resemble the ground truth, despite the loss being relatively low. The model seems to be predicting only the background, failing to capture the lesion area.

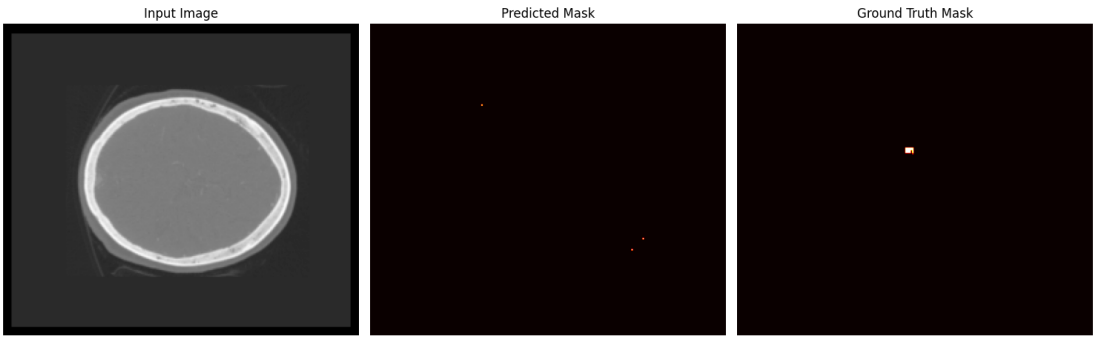
Upon further analysis, we found that the BCE loss function is not well-suited for this problem. This is due to the fact that the positive class (the lesion) is often much smaller than the background. Using BCE encourages the model to focus on the background, leading to:

- False negatives, where small lesions are barely penalized
- Uneven learning, where the model learns to detect background features rather than the lesion itself

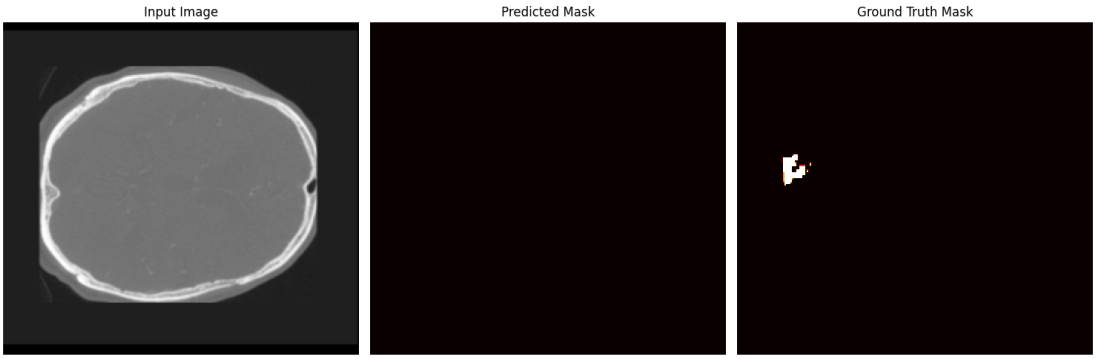
In contrast, the Dice loss function is more appropriate for this task, as it directly maximizes the spatial overlap between the predicted mask and the ground truth, which is crucial when dealing with imbalanced datasets like this.

4.4 Dice Loss Function

The Dice loss function is commonly used in medical image segmentation tasks due to its effectiveness in handling class imbalance, which is typical in medical imaging where the region of interest (e.g., lesions) often occupies only a small portion of the image.



(a) First example of a CT scan slice with corresponding prediction and ground truth mask. The prediction shows no correspondence with the actual lesion.



(b) Second example showing a different CT scan slice. Here too, the model fails to detect the lesion, predicting only background.

Figure 4.1: Comparison between input images (left), predicted masks (middle) and ground truth masks (right) for two different examples. The black prediction masks show that the model mainly predicts the background and fails to detect the lesions.

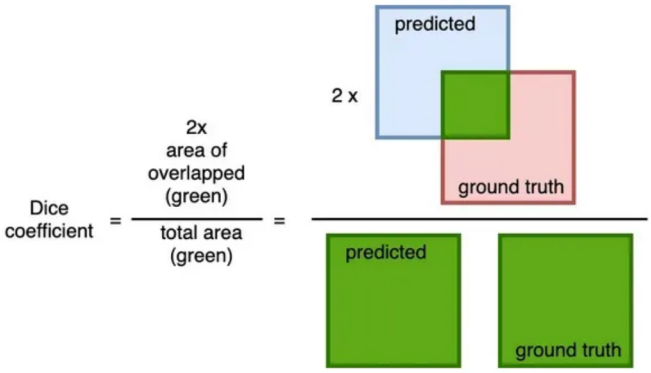


Figure 4.2: Visual explanation of the Dice coefficient calculation. The Dice coefficient is computed as the ratio of twice the area of overlap (green) to the total area of both the predicted segmentation and ground truth. Image adapted from [Nev23].

The Dice loss is derived from the Dice coefficient (also known as F1 score), which measures the overlap between the predicted segmentation and the ground truth. The Dice coefficient is defined as:

$$\text{Dice} = \frac{2|X \cap Y|}{|X| + |Y|} \tag{4.1}$$

where X is the predicted set and Y is the ground truth set.

List of Figures

List of Tables

Bibliography

[Nev23] Neville. *Dice Loss In Medical Image Segmentation*. July 2023. URL: https://cvinvolution.medium.com/dice-loss-in-medical-image-segmentation-d0e476eb486 (visited on 04/17/2025).