

PROJECT REPORT

Machine Learning - UML501

The Liver Tumour Segmentation

Somil Kumar 102203221

Pratyush Sharma 102203958

Dilpreet Singh 102203004

Submitted To:

DR. ARCHANA SINGH



THAPAR INSTITUTE
OF ENGINEERING & TECHNOLOGY
(Deemed to be University)

Department of Computer Science & Engineering

Introduction

The **Liver Tumor Segmentation Benchmark (LiTS)** is a large-scale dataset and competition designed to improve and evaluate the performance of machine learning models in accurately segmenting liver and liver tumors from **computed tomography (CT) scans**. It was created to encourage research into developing automated solutions for liver tumor diagnosis and treatment planning.

The challenge involves:

1. **Dataset:** A collection of 201 CT volumes annotated by expert radiologists, covering a wide range of liver tumor types and conditions.
2. **Task:** Automated segmentation of liver and liver tumors using advanced algorithms, primarily deep learning.
3. **Benchmarking:** Providing standardized evaluation metrics (like Dice score and surface distance) to rank the performance of various algorithms.

Why is it Important?

Liver cancer is one of the most fatal cancers globally. Accurate detection and segmentation of liver tumors are critical for:

- **Diagnosis:** Identifying abnormalities in the liver for early detection of cancer.
- **Treatment Planning:** Determining the location, size, and spread of tumors for therapies like surgery, radiotherapy, or ablation.
- **Monitoring Progress:** Assessing the response to treatment by comparing tumor sizes over time.

Manually segmenting liver tumors from CT scans is time-consuming, prone to variability between radiologists, and challenging due to:

- **Tumor diversity:** Wide variations in tumor size, shape, and contrast.
- **Background complexity:** Similarities between tumor textures and surrounding liver tissue, especially in cases with liver diseases like cirrhosis.

LiTS helps researchers develop reliable, automated tools to address these challenges, making it a significant step forward in medical imaging and cancer care.

Literature Review

The Liver Tumor Segmentation Benchmark (LiTS)

Authors: Patrick Bilic, Patrick Christ, Hongwei Bran Li, Eugene Vorontsov, et al.

This paper reports on methods, results, and challenges faced by participants in the LiTS competitions. Below is a summary of the reported methods, parameters, and performance metrics for liver and liver tumor segmentation.

Methods

1. Method: Cascaded U-Net (Han et al.)

- **Architecture:**
A **2.5D residual U-Net** using stacked CT slices as input. It performs liver segmentation first, followed by tumor segmentation.
- **Data Preprocessing:**
HU value clipping to a range of **[-200, 200]** to normalize intensity, along with flipping and cropping for data augmentation.
- **Post-processing:**
False positives removed by keeping only connected components with a probability > 0.8 .
- **Key Parameters and Metrics:**
 - **Accuracy:** Achieved a Dice score of **0.963** for liver segmentation and **0.674** for tumor segmentation (ISBI 2017).
 - **Precision and Recall:** Precision = **35.4%**, Recall = **45.8%** for lesion detection.
 - **Time:** Model training and evaluation time was reasonable, with segmentation per scan taking ~2 minutes.

2. Method: H-DenseUNet (Li et al.)

- **Architecture:**
A **hybrid DenseNet** combining 2D and 3D convolutions for better feature extraction from spatial and contextual information.
- **Data Preprocessing:**
HU clipping to **[-200, 250]**, rotation, flipping, and scaling for augmentation.

- **Post-processing:**
Used a connected component filter to remove false positives.
- **Key Parameters and Metrics:**
 - **Accuracy:** Dice score of **0.702** for tumor segmentation (MICCAI 2017).
 - **Precision and Recall:** Precision = **14.8%**, Recall = **47.9%** for lesion detection.
 - **Time:** More computationally intensive due to hybrid 2D-3D processing.

3. Method: nn-UNet (Isensee et al.)

- **Architecture:**
An **auto-configuration U-Net** that adapts to the dataset's specific characteristics. It excels by automatically adjusting preprocessing, augmentation, and network depth.
- **Data Preprocessing:**
HU clipping, intensity normalization, and advanced geometric augmentations.
- **Post-processing:**
Enforced tumor location within the liver using liver masks.
- **Key Parameters and Metrics:**
 - **Accuracy:** Dice score of **0.739** for tumor segmentation (MICCAI 2018).
 - **Precision and Recall:** Precision = **55.4%**, Recall = **55.4%** (balanced for detection).
 - **Time:** Higher computational cost, suitable for high-resource environments.

Comparison

Method	Dice Score	Precision	Recall	Processing Time
Cascaded U-Net	Liver: 0.963 Tumor: 0.674	35.4%	45.8%	~2 minutes/scan
H-DenseUNet	Liver: 0.961 Tumor: 0.702	14.8%	47.9%	High (2D-3D hybrid)
nn-UNet	Liver: 0.962 Tumor: 0.739	55.4%	55.4%	Higher computational cost

Methodology

1. U-Net (Deep Learning Model)

- **Description:** A fully convolutional neural network widely used for biomedical image segmentation.
 - **Implementation:** Using MONAI's U-Net implementation, pre-trained with standard configurations for 3D CT volumes.
 - **Rationale:** Ideal for segmenting organs like the liver due to its encoder-decoder architecture with skip connections that preserve spatial information.
-

2. Random Forest (Machine Learning Model)

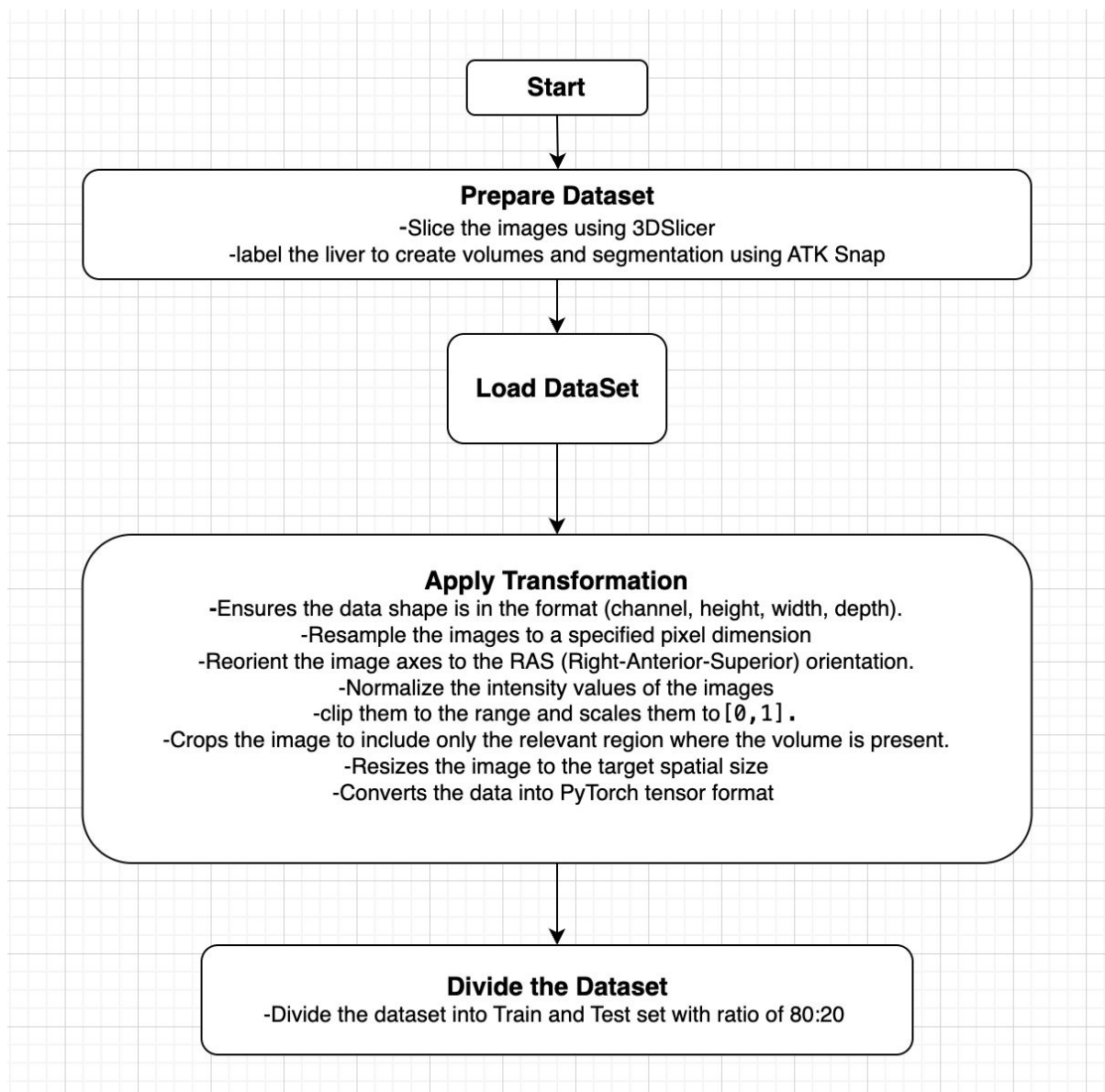
- **Description:** A tree-based ensemble model that uses random sampling and averaging to perform robust classification.
 - **Implementation:** Extract features from CT images (e.g., intensity, texture, edge details) and use them to train a Random Forest classifier for segmenting liver regions.
 - **Rationale:** Useful as a lightweight model for initial liver detection, requiring less computational power than deep learning.
-

3. Support Vector Machine (SVM) (Machine Learning Model)

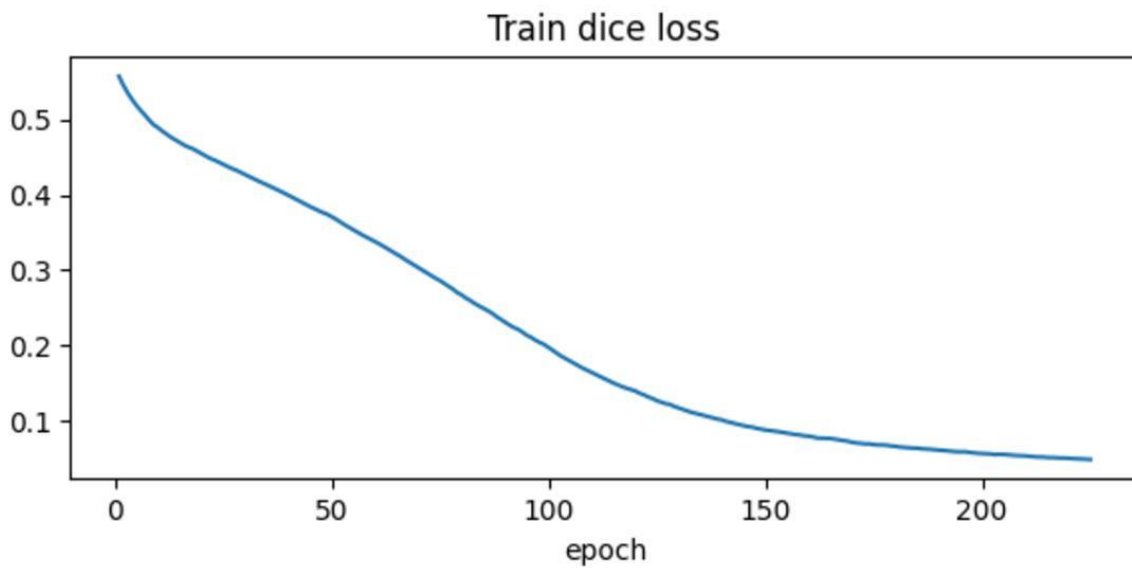
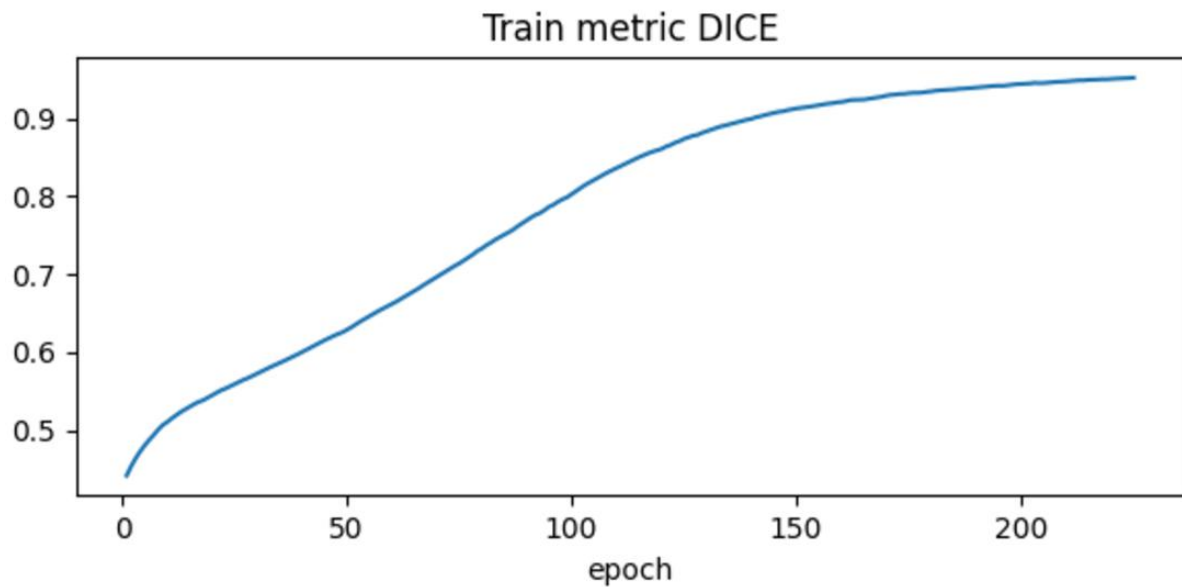
- **Description:** A supervised learning algorithm that finds the hyperplane separating different classes in the feature space.
- **Implementation:** Use CT image features (similar to Random Forest) as input for SVM classification, focusing on separating tumor and healthy liver regions.
- **Rationale:** Effective for binary classification problems, making it a good fit for distinguishing tumor vs. non-tumor regions.

Preprocessing

This flowchart shows the steps involved in preparing and processing a dataset for liver tumor segmentation. First, the images are sliced and labeled using 3D Slicer and ATK Snap. Then, the dataset is loaded, and various transformations are applied to ensure the data is in the correct format, resized, and normalized. The images are cropped to focus on the relevant region and converted into a format suitable for PyTorch. Finally, the dataset is divided into training and testing sets with an 80:20 ratio.

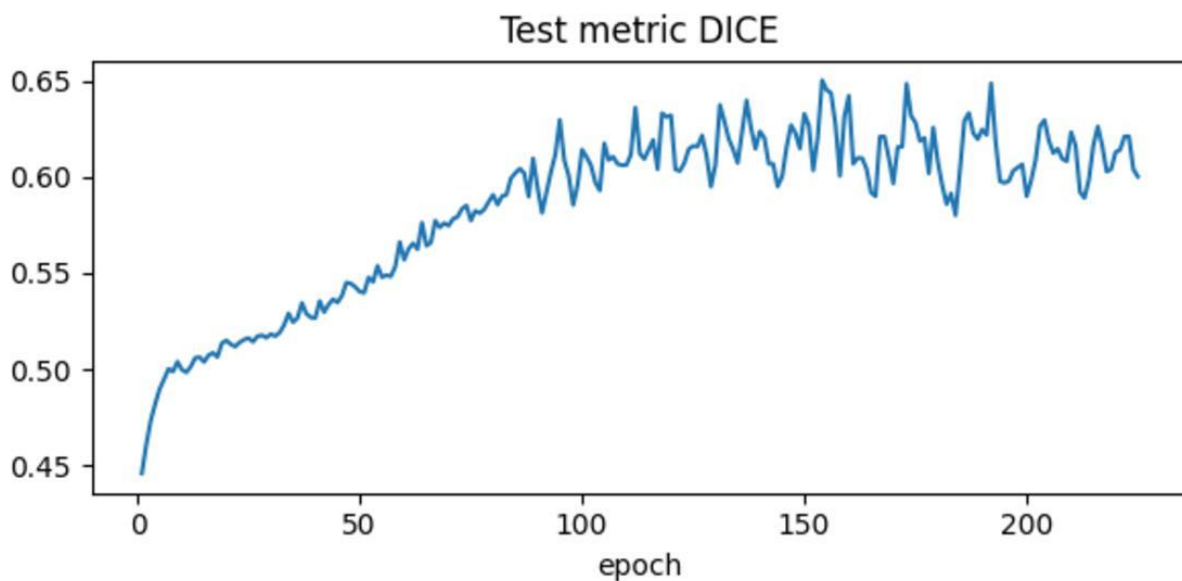
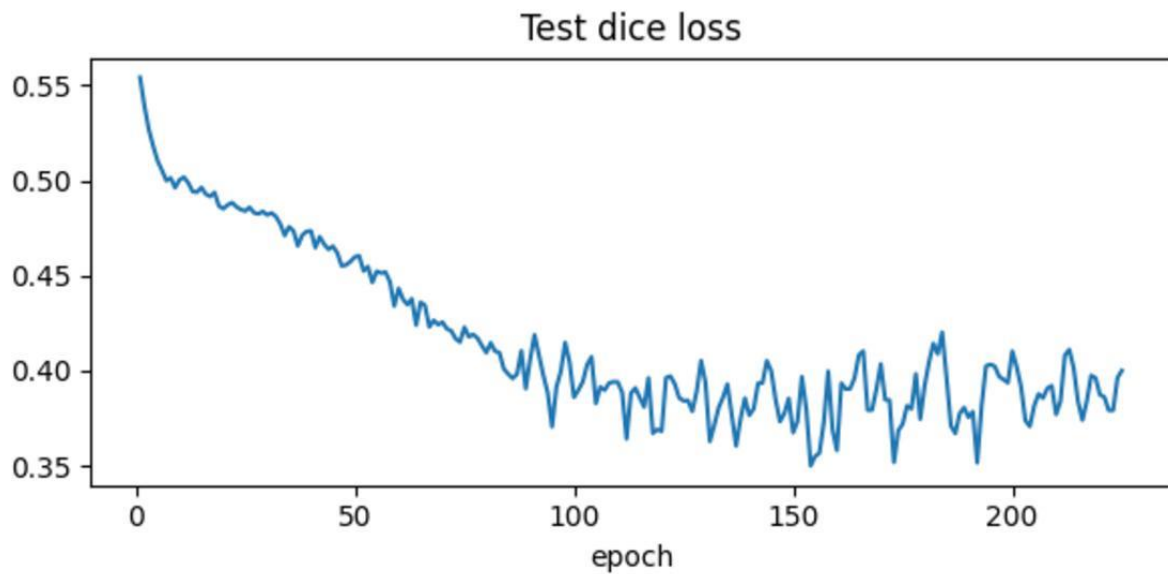


Training



```
Epoch_loss: 0.0483
Epoch_metric: 0.9517
test_loss_epoch: 0.4000
test_dice_epoch: 0.6000
current epoch: 225 current mean dice: 0.6905
best mean dice: 0.6501 at epoch: 154
train completed, best_metric: 0.6501 at epoch: 154
```

Testing



Confusion Matrix (Overall):

```
[[10183590.    24984.]
```

```
[ 221079.    56107.]]
```

Accuracy: 0.9765

Precision: 0.6919

Recall: 0.2024

Dice Metric: 0.3132

Dice Loss: 0.6868

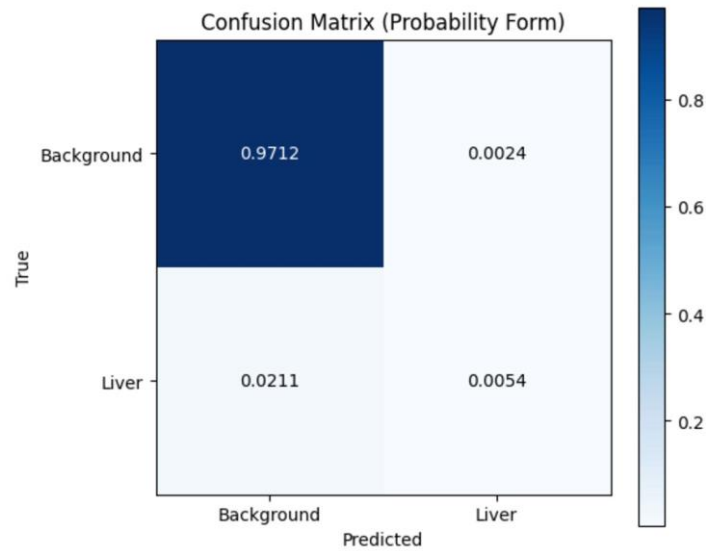
print(confusion_matrix(ProbabilityForm

Confusion Matrix (Probability Form):

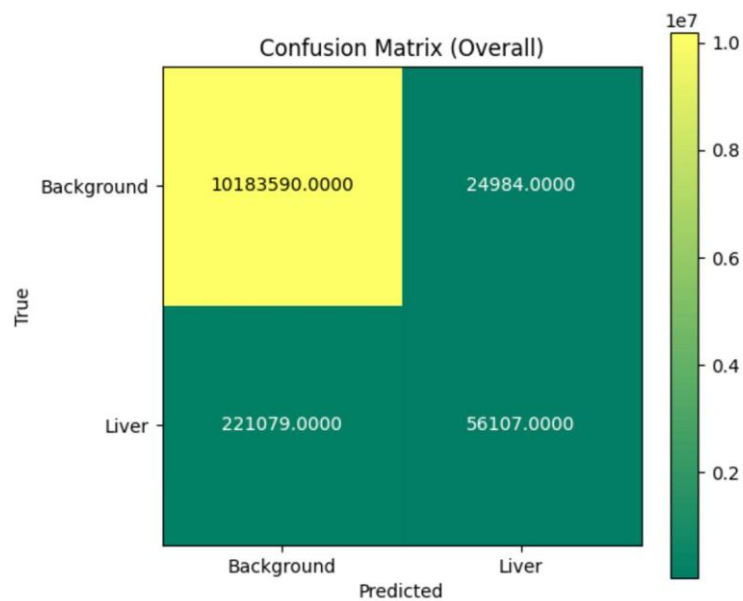
`[[0.97118282 0.00238266]`

`[0.02108374 0.00535078]]`

import seaborn as sns



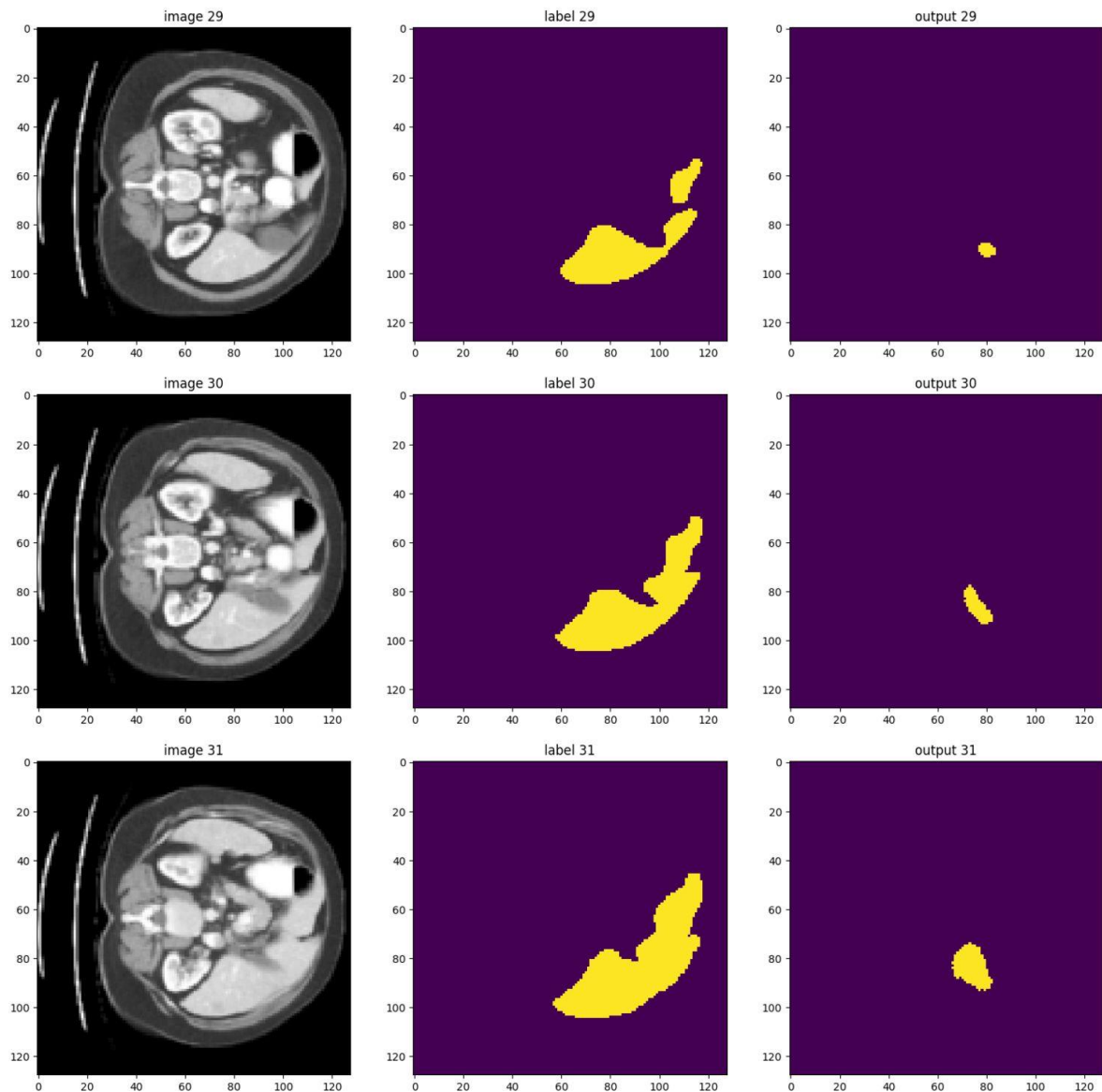
Probability of True Positive (TP): 0.9712
Probability of False Positive (FP): 0.0024
Probability of False Negative (FN): 0.0211
Probability of True Negative (TN): 0.0054



Number of True Positive (TN): 10183590.0000
Number of False Positive (FP): 24984.0000
Number of False Negative (FN): 221079.0000
Number of True Negative (TP): 56107.0000

Result

In the image, the "label" represents the liver, and the "output" corresponds to the tumor segmentation.



Links

Code : <https://github.com/Pratyushs411/LiverSegamentation>

Reference:

<https://www.sciencedirect.com/science/article/pii/S1361841522003085>