Report Medical Images Processing

lyes.boudia

May 2025

1 Introduction

This project is a complete pipeline for loading, analyzing, and visualizing 3D medical CT scans, with a special focus on segmenting and highlighting liver and tumor regions. It uses real medical imaging data in the DICOM format, which is a standard used in hospitals and clinics for storing CT, MRI, and other medical images. The dataset comprises a series of CT scan slices and two corresponding segmentation files that delineate the liver and tumor areas. The core objectives include reconstructing a 3D volume from these 2D slices, aligning and displaying the liver and tumor segmentations over the CT images, and analyzing voxel dimensions to understand the real-world size of each pixel. Furthermore, the project aims to create projections and animations for better comprehension of internal structures and tumor positioning, coregister the input to a reference image by implementing all necessary steps, and finally, visualize the liver region in the input image space to assess the algorithm's correctness, both numerically and visually.

2 DICOM loading and visualization

The pipeline begins with the loading of a DICOM series containing axial CT slices. Using the pydicom library, the slices are sorted based on the ImagePositionPatient DICOM attribute, which encodes the 3D spatial location of each image in patient space. This step ensures that the volumetric reconstruction respects the anatomical order from inferior to superior. The pixel data from each slice is stacked into a 3D NumPy array to form the complete CT volume. Spatial resolution is preserved by extracting voxel spacing information from the PixelSpacing and SliceThickness tags in the metadata.

For segmentation data, we load DICOM-SEG files using highdicom, which allows us to parse standard-compliant DICOM segmentation objects. These files contain multiple labeled regions corresponding to liver and tumor masks. Each segmentation frame is linked to a specific CT slice via its Z-coordinate (from ImagePositionPatient), enabling us to align the segmentation masks with the CT volume. We map each frame to its corresponding slice in the CT volume and populate 3D binary arrays representing the liver and tumor masks.

To produce visual overlays, we first window the CT image using a center of 150 and width of 900 to enhance soft tissue contrast. The liver and tumor masks are then rendered using color blending: green for the liver and red for the tumor. Alpha blending is used to visually merge the masks with the underlying CT, with the tumor mask rendered more opaquely to ensure it is visually distinguishable from the liver. In overlapping regions, the tumor mask takes precedence to emphasize pathological structures over normal anatomy.

We then generate various visual representations of the 3D data. Multi-planar reformats (MPRs) are produced in both sagittal and coronal planes by slicing or projecting the volume orthogonally. For each plane, three types of views are computed: (1) median slice, showing the anatomical center of the volume; (2) maximum intensity projection (MIP), which projects the highest intensity values along the viewing axis, emphasizing dense structures like bones and contrast-enhanced regions; and (3) average intensity projection (AIP), which provides a smoother view of soft tissues by averaging across slices. These views help clinicians understand both the spatial relationships and intensity distributions in different anatomical planes.

Finally, we produce a dynamic visualization by rotating the CT and fused segmentation volume around the axial (Z) axis. This is implemented by generating sagittal MIPs at 10-degree intervals over a full 360-degree rotation. Each projection is rendered as a frame, and the full sequence is compiled into an animated GIF. This rotating view provides a comprehensive understanding of 3D spatial relationships, especially useful for evaluating tumor position relative to liver anatomy.

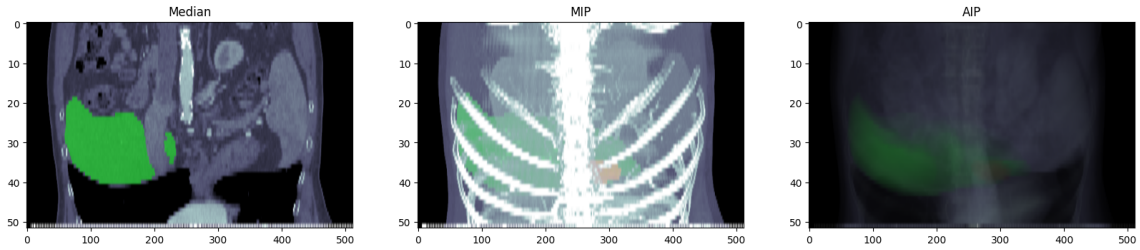

Figure 1: Coronal planes

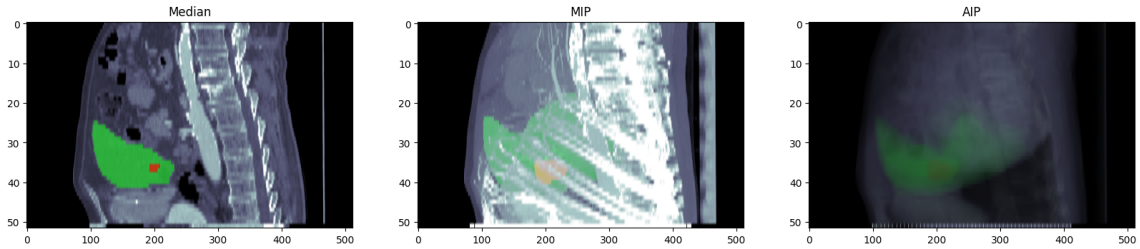

Figure 2: sagittal planes

3 3D Rigid Coregistration

This pipeline starts by setting paths to the dataset directories containing two DICOM image series: a reference volume and an input volume. These series are loaded into 3D arrays along with their metadata, which include important spatial information like slice thickness and pixel spacing. The original volumes are inspected visually to check quality and structure. Then, the pipeline extracts

and prints detailed pixel dimensions for both volumes, capturing the physical size of each voxel in the Z (slice thickness), Y, and X directions. This is crucial for accurate spatial understanding.

Next, the pipeline calculates the gradient magnitude volumes for both datasets, which highlight edges and structures by measuring changes in intensity. Comparing these gradient volumes with a Sum of Squared Differences (SSD) metric and Mutual Information (MI) gives a quantitative baseline of how similar or different the volumes are before any registration.

To prepare for efficient processing, both volumes are normalized to a consistent intensity scale, ensuring comparability during visualization and registration. The volumes are then downsampled in the in-plane directions (Y and X) by a factor of two, reducing data size while maintaining slice thickness resolution. This downsampling speeds up subsequent computations without sacrificing too much detail.

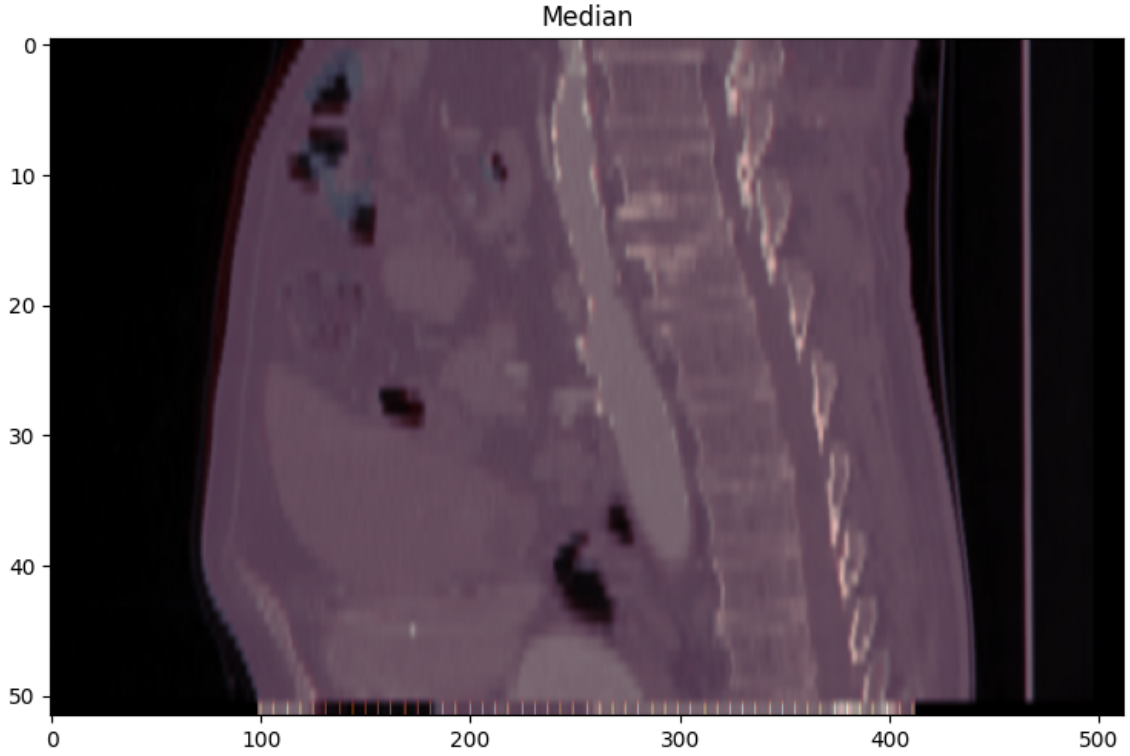

Figure 3: compares the reference volume and the input volume before coregistration using sagittal planes

Following downsampling, the volumes are resampled to achieve isotropic voxel spacing—meaning voxels have equal physical dimensions along all three axes—set here to 1 mm^3. This standardization is important for registration algorithms, which perform better with uniform voxel sizes. The metadata is updated accordingly to reflect these changes.

With preprocessing complete, a gradient-based rigid coregistration algorithm is applied. This

algorithm optimizes a set of transformation parameters, including translations along the three axes, rotation defined by an axis and angle, and scaling factors, to best align the input volume to the reference. The optimization progress and final results (cost function, number of iterations, optimized parameters) are displayed for transparency.

The optimized parameters are converted from normalized scales into physical units—translations in millimeters, rotation angle in degrees, and scaling factors close to 1—to make them interpretable in the real world. The rotation axis is normalized and displayed to understand the direction of rotation.

Finally, the pipeline applies the optimized rigid transformation to the input volume, producing a registered volume aligned to the reference. Visual overlays of sagittal, coronal, and axial median slices from both volumes are plotted using different colormaps and transparencies to qualitatively assess registration success. These visualizations help confirm how well the input data has been spatially aligned after the entire pipeline.

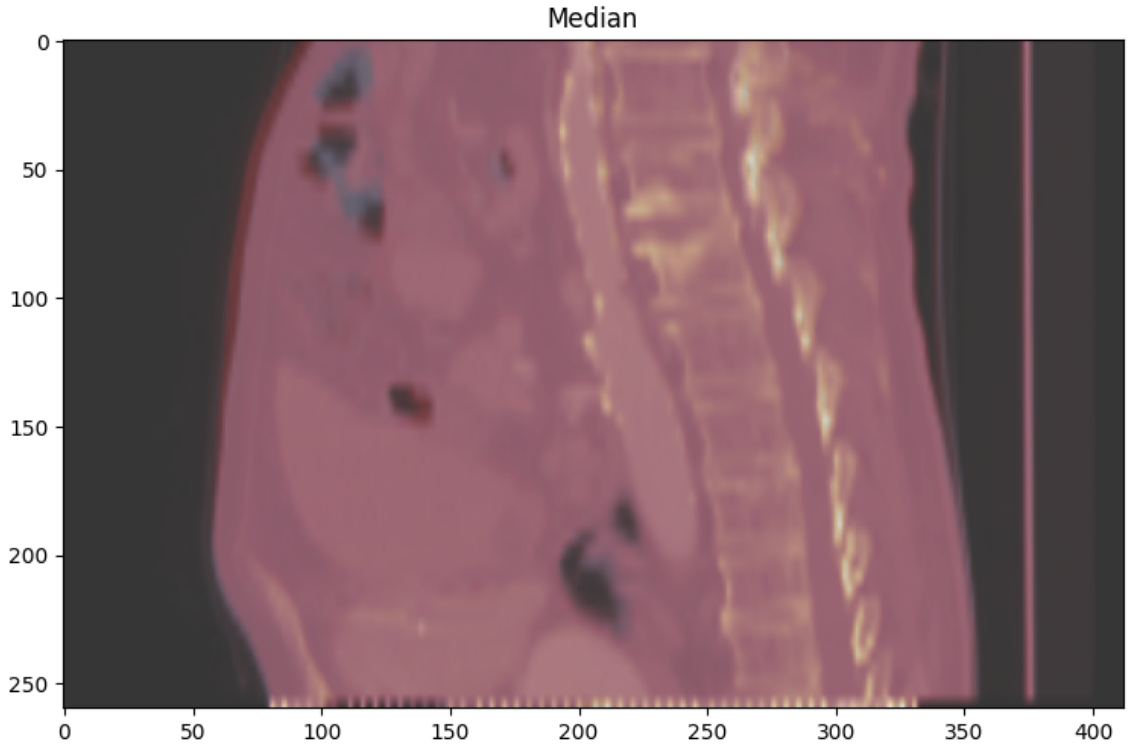

Figure 4: compares the reference volume and the input volume transformed after coregistration using sagittal planes

After completing volume registration, the pipeline quantitatively assesses how well the input volume aligns with the reference volume using multiple complementary metrics. First, it calculates the gradient magnitude of both the resampled reference volume and the transformed (registered)

input volume. Gradient magnitudes emphasize structural edges, so comparing them via Sum of Squared Differences (SSD) provides insight into how closely anatomical boundaries match. Mutual Information (MI), a statistical measure capturing intensity distribution similarities, is also computed to evaluate overall intensity correspondence between the volumes.

Next, the pipeline loads a manual liver segmentation mask from the reference image space, which defines the ground truth liver region. This mask undergoes the same downsampling and isotropic resampling steps as the reference volume, ensuring spatial alignment with the processed data. To compare segmentation masks in the input volume's coordinate system, the pipeline applies the inverse of the previously optimized rigid transformation to the resampled liver mask, effectively mapping the reference liver segmentation into the input volume space.

Visual overlays of the transformed liver mask on the input volume are created across sagittal, coronal, and axial median planes using transparent color maps to qualitatively verify alignment of the liver region post-registration.

For quantitative overlap evaluation, the pipeline loads the ground truth liver segmentation from the input volume space, resamples it identically to the input volume, and pads it to ensure shape compatibility with the transformed liver mask. Both masks are binarized and verified to have matching dimensions, which is critical for voxel-wise comparison.

The key overlap metrics calculated are the Dice Coefficient and Jaccard Index. Dice measures the harmonic mean of precision and recall between two masks, indicating volumetric similarity, while Jaccard measures intersection over union. The pipeline reports high values (Dice 0.97, Jaccard 0.94), reflecting excellent agreement between the registered segmentation and ground truth.

Additionally, the pipeline computes Sensitivity (Recall) and Specificity. Sensitivity quantifies the proportion of true liver voxels correctly identified by the transformed mask, and specificity quantifies the proportion of non-liver voxels correctly excluded. Near-perfect specificity (0.999) and high sensitivity (0.97) further confirm precise liver localization after registration.

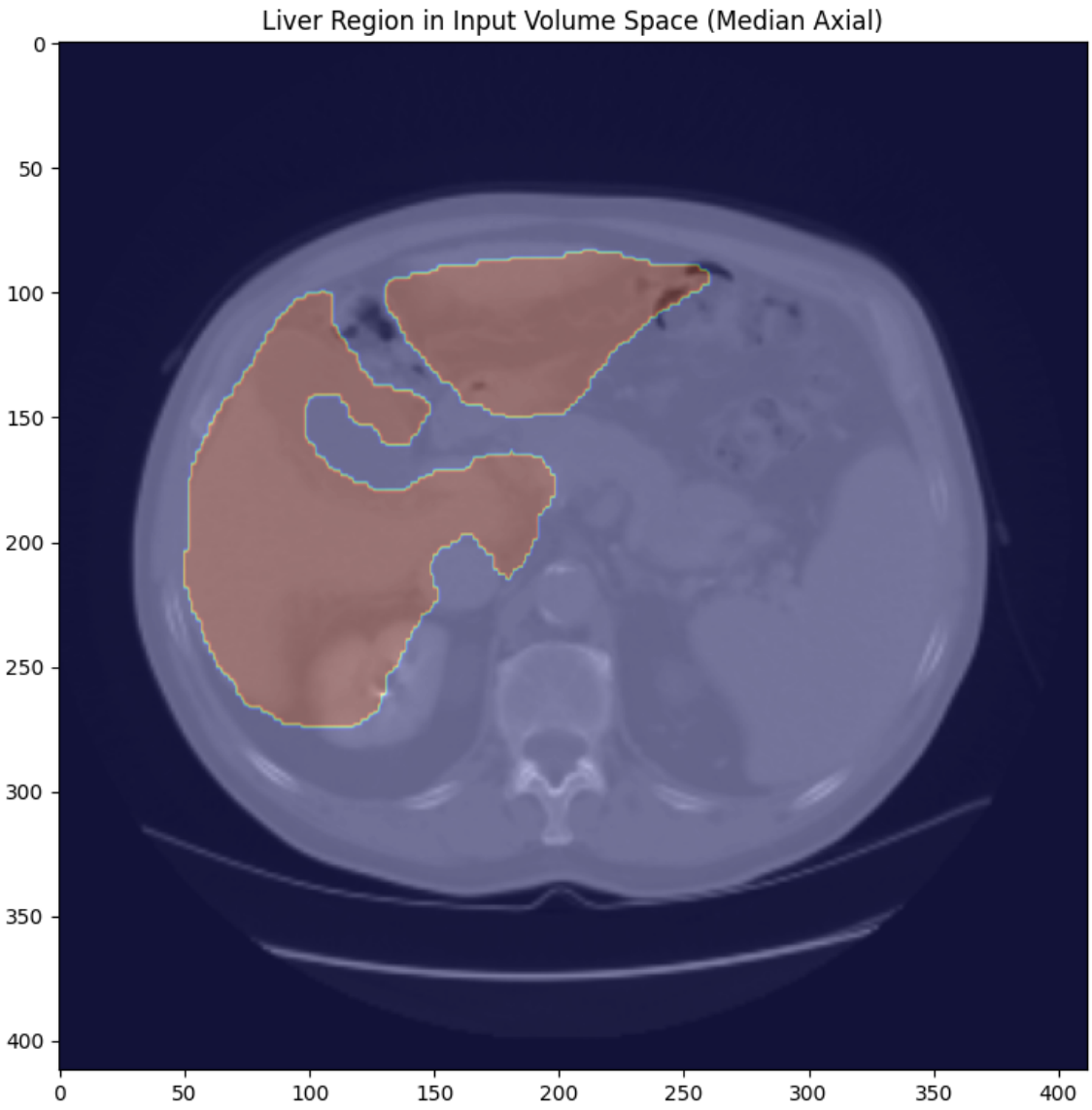

Figure 5: Liver Region in Input Volume Space (Median Axial)

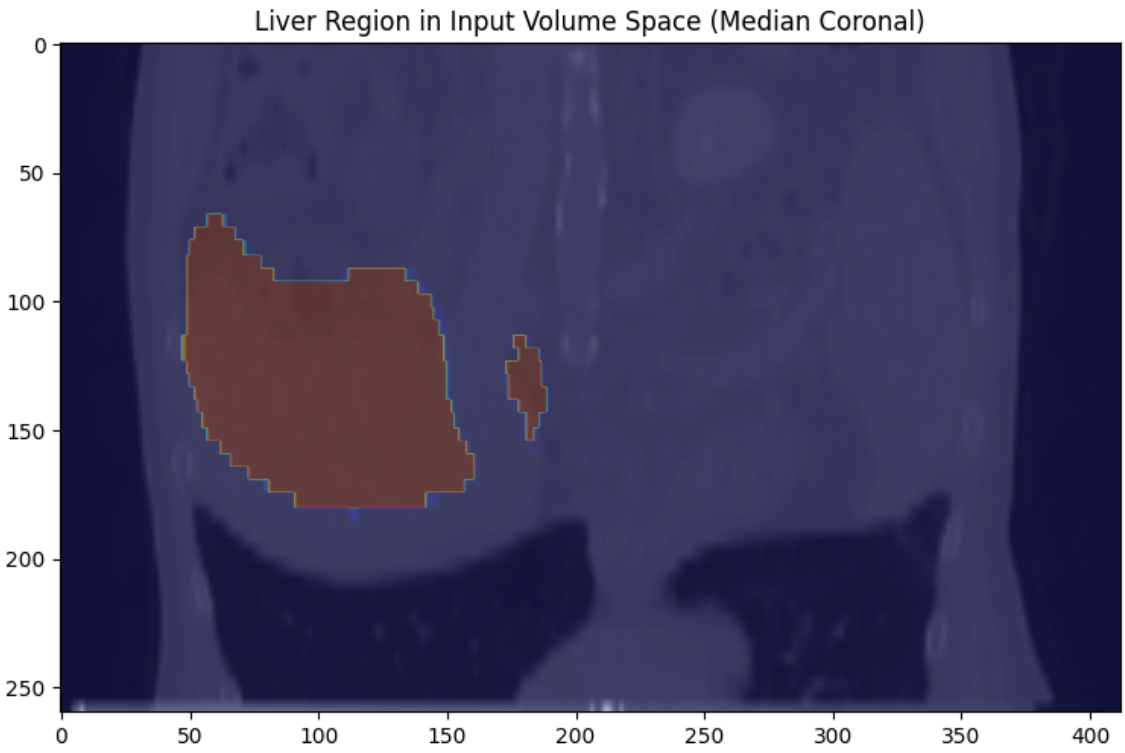

Figure 6: Liver Region in Input Volume Space (Median coronal)

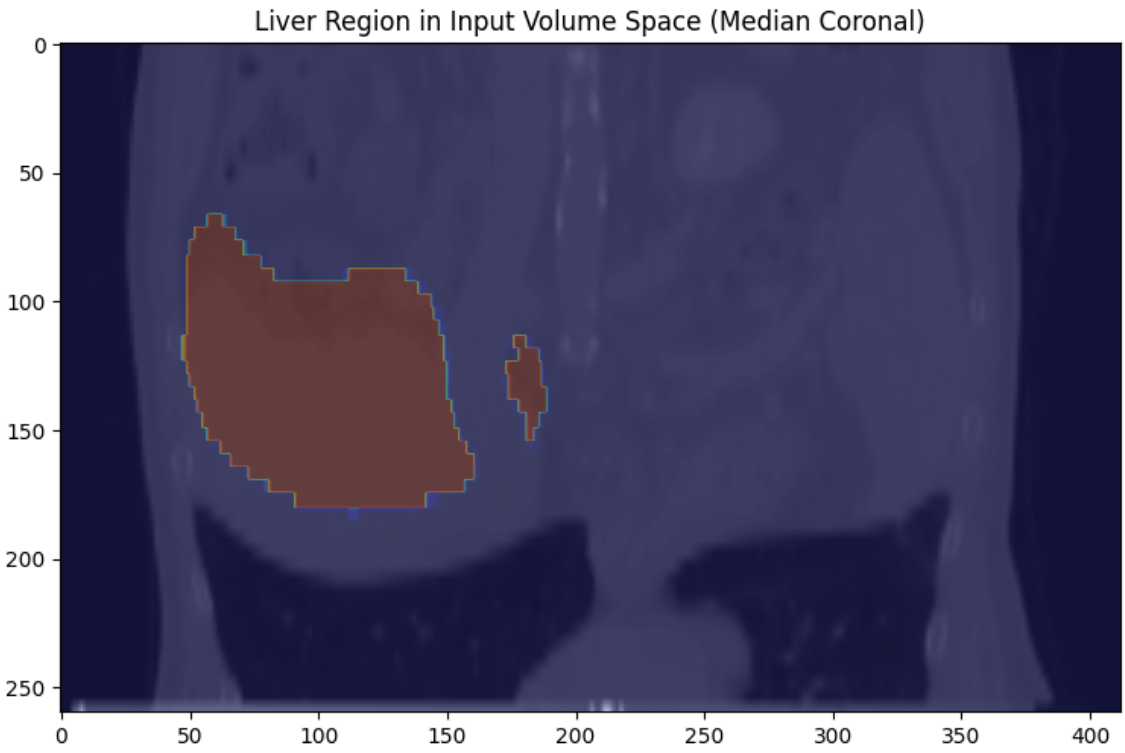

Figure 7: Liver Region in Input Volume Space (Median sagittal)