Multimodal Imaging for Cancer Detection: PET-CT Reconstruction & AI Segmentation

Lewis McConkey
Cambridge University
Word count: 2997

April 3, 2025

Abstract

This report highlights the importance of choosing the appropriate reconstruction algorithm for both CT and PET scans by testing out various iterative and non-iterative methods (OSEM, MLEM etc.). We present different reconstructed images, comparing them at various iterations before highlighting that OSSART is best for CT reconstruction while arguing that OSEM reconstructs higher quality images from PET scans after attenuation correcting. Initial results from a combined image acquired from six coil receivers (for 3D MRI k-space data from a knee scan) show the impact of the ringing artefact. After comparison of several denoising methods however wavelet denoising is shown to improve the image quality while methods like PRI-NL-PCA and DNN's have been part of recent breakthroughs in the denoising space. Finally we discuss methods for segmenting tumours and feature extraction along with highlighting that a CNN with minimal preprocessing has also been shown to reach the performance close to human observers on brain images.

1 Introduction

Medical imaging is essential for tumor detection and classification in MRI and CT scans but often requires preprocessing to reduce noise and artifacts. This project explores reconstruction algorithms for CT and PET scans, denoises MRI data, and extracts histogram-based features (intensity, MAD, uniformity) to analyze tumor characteristics. We develop a segmentation function and use radiomic features to classify tumors as malignant or benign.

2 PET-CT Reconstruction

For CT we are given the sinogram, the dark fields (scan to measure the background noise without X-rays) and the flat fields (Captured by exposing the detector to a uniform X-ray beam (without a patient)). For PET we are given the sinogram and a detector gain calibration sinogram (compute the gain of each of them given the same input signal). We display plots of all of these in the accompanying code but in this report we just show the two sinograms:

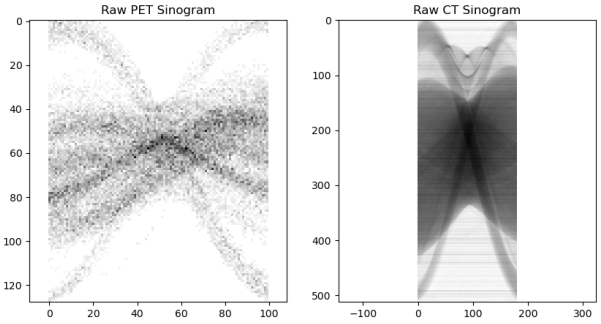

Figure 1: PET and CT Sinograms

We can see that the CT sinogram has been acquired around 180 degrees with one measurement per degree and also see the background noise in the image (mainly the black horizontal lines across the image). With the PET sinogram we can see that the scan produces dark spots where there should be bright spots and vice versa.

2.1 Cleaning Sinograms

First cleaning the CT sinogram we remember that we have the dark field image and we need to subtract this away from the sinogram to remove the unwanted background noise produced by the detector. Parts of the detector are more/less sensitive to X-rays so we divide by the flat field to normalise these differences but the dark fields contribute to this too so we need to take this away. This can be seen in the accompanying code and with the corrected CT sinogram plotted below:

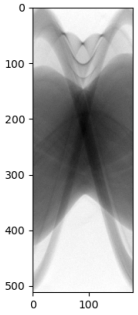

Figure 2: Corrected CT sinogram

We can see that all the noise has been removed and the image looks improved and cleaner. Now with the PET sinogram we want to ensure consistent sensitivity across all PET detector elements. We can see from the original sinogram that some areas appear darker and some appear brighter than they should be and hence we divide by the PET calibration image to correct for these detector response variations. This can be seen in the accompanying code and with the corrected PET sinogram below:

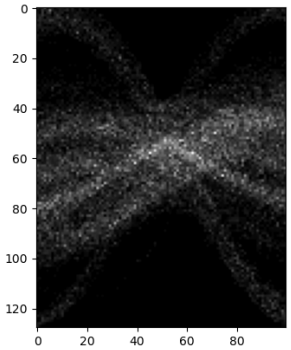

Figure 3: Corrected PET sinogram

2.2 Reconstruct CT image

We first reconstruct the CT image using filtered backprojection (FBP) which is a direct reconstruction method that uses a fllter and then back projects (by putting each value from the detector into the image). We use the following integral for FBP:

$$f_{fbp}(x,y) = \int q_\theta(xcos(\theta) + ysin(\theta))d\theta$$

$$\text{where } q_\theta(t) = \int P_\theta(\omega)|\omega|e^{2\pi i\omega t}d\omega$$

x and y are the points we want to reconstruct in the image and the integral is taken over all projection angles by summing up the contributions of the projections taken at different angles. The function $q_\theta(t)$ represents the filtered projection data for each angle. This equation represents the radon transform so when we want to reconstruct the image using FBP we just use the inverse radon transform as seen in the code. This is the image we have reconstructed:

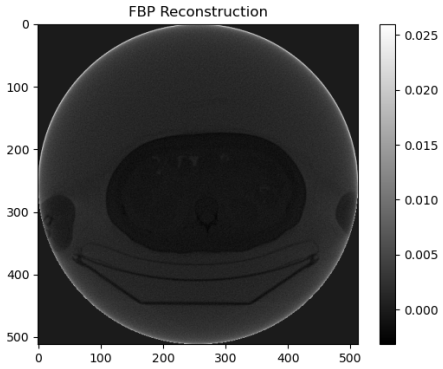

Figure 4: Image reconstruction using FBP

We next use OS-SART which is a stochastic version of SART (the first iterative algorithm proposed in CT). This means OS-SART is given by:

$$x^{K+1} = x^k + \gamma A_i^T (A_i x^k - b)$$

We use iradon_sart in python with a for loop for the iterations to give OS-SART. We initialise the algorithm by choosing $K = 10$ and $\gamma = 0.2$ and aim to improve the performance by hyper-parameter tuning these. When we use $K = 50$ and $= 0.2$ we get an image that is starting to improve on FBP, which can be seen here:

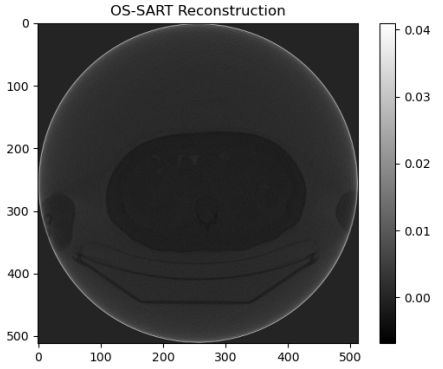

Figure 5: Image reconstruction using OS-SART

The brighter regions are far more distinct and sharper with OS-SART whereas with the FBP from before the brighter regions have kinks in them and don't have sharp edges like a perfect circle. This is something that we observed with all varying chooses of hyper-parameters from using OS-SART indicating this is something that generally OS-SART just improves over FBP. The blurriness is still inherent in both but we can observe that OS-SART starts showing less noise in the image when you decrease iterations to around 50. When we decrease iterations any lower than around 10 and decrease γ to less than around 0.2 we observe that OS-SART doesn't really improve or decrease in quality and keep the optimal model with those as our hyper-parameters. This can be seen in the accompanying code to show a reconstructed image that is more well defined, sharper and with less noise than the FBP reconstructed image in this report. The following table shows the performance of SIRT, OS-SART and FBP:

Algorithm	Iterations	time (s)
OS-SART	10	8-14
SIRT	50	132
FBP	N/A	0.25-0.5

From this we view that FBP is much faster than the others because it is not an iterative algorithm but it is known to be sensitive to noise and artifacts as can be seen in our reconstruction and in [1]. OS-SART is more than 10 times slower than FBP but presents better performance which is also shown in [1] where they reconstruct images from a Digital Breast Tomosynthesis (DBT) system. OS-SART requires much less iterations (10 compared to 50 used in CT practical) to get as good an image reconstruction as SIRT and scales linearly with time so still takes a considerably amount less in time than SIRT does. This is because in OS-SART we are only using a subset of A (only forward/backprojecting some angles) but in the CT practical (using SIRT) we processed all the projections at each iteration making it slower.

2.3 PET Attenuation Correction

PET imaging suffers from photon attenuation, where emitted photons interact with tissues and fail to reach the detector due to either absorption or scattering. This could lead to poor image quality and central regions of the image to appear artificially darker because photons traveling through more tissue are more attenuated. The choice between attenuated and non-attenuated correction (NAC) is a huge research area with conflicting results as [2] shows that non attenuated correction provides better visibility in 41.4% of lesions however in [3] attenuation correction (AC) improves lesion detection for DH coincidence imaging. This highlights that even as PET/CT use becomes more

widespread, evaluation of both NAC and AC images should remain an integral part of image interpretation. With this in mind we move on with CT based AC by resizing the OSSART reconstructed CT image using the pixel sizes. We produce the attenuation sinogram for this and resize it (to match the size of the PET image) , which allows us to AC the PET sinogram using the Beer-Lambert sinogram. This is our final image:

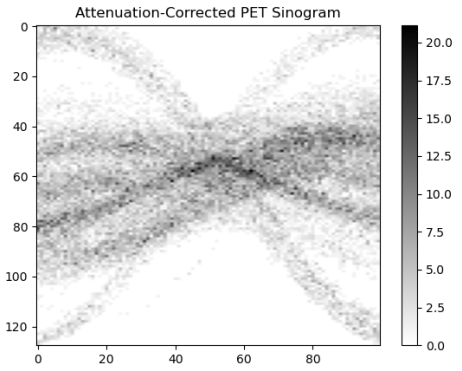

Figure 6: Attenuation-Corrected PET sinogram

The image shows a well-applied attenuation correction although further evaluation could provide more insights such as measuring standardised uptake values in key regions like in [2].

2.4 Reconstruct PET Image

FBP has become an outdated image reconstruction technique in new-generation PET scanners [4] which can be observed with the poor reconstruction that FBP produces in this case:

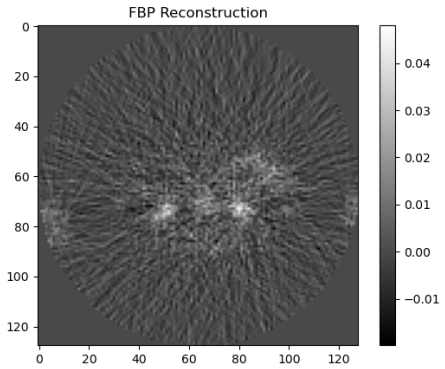

Figure 7: PET image reconstruction using FBP

This happens because FBP performance rapidly decays with increase of noise and reduction of projections. PET scans collect fewer photons (they suffer from photon attenuation and scatter) leading to high noise levels and hence iterative methods are preferred. We test out two iterative methods to reconstruct the PET image: OSEM and MLEM. The 2 reconstructions are shown below:

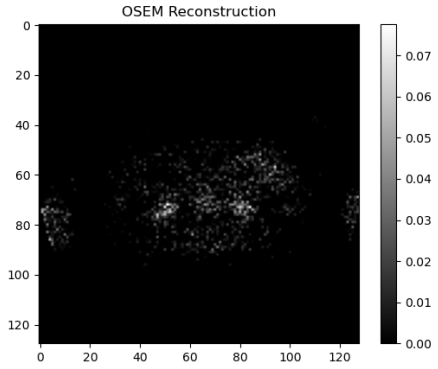

Figure 8: PET image reconstruction using OSEM (20 iterations)

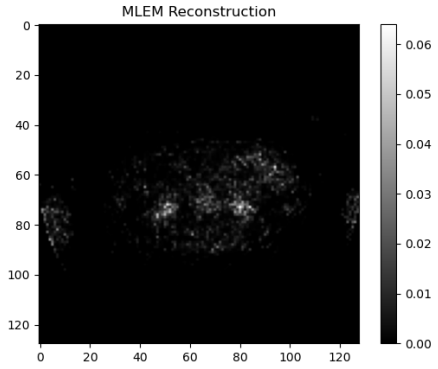

Figure 9: PET image reconstruction using MLEM (50 iterations)

We can observe that OSEM converges to a better result and higher quality image in less iterations with OSEM taking 20 iterations to reach a better reconstruction than MLEM does in 50 iterations. This is because each subset processes part of the data so the image updates more often compared to MLEM, which waits for a full pass over the data before updating. OSEM with 4 subsets is around 4 times faster per iteration than MLEM, with 8 subsets is 8 times faster and so on [5] which highlights why OSEM is the standard used in PET reconstruction.

2.5 Theory

1. We first overlay the CT and PET scans and produce this heatmap:

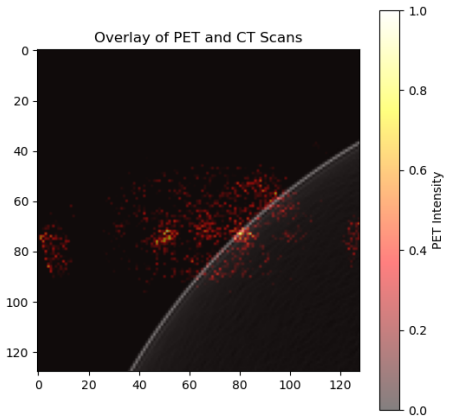

Figure 10: PET and CT scans overlayed

We normalised the images so they are on a similar intensity scale and used transparency to overlay the PET scan on top of the CT scan. Fusing these two modalities together by overlaying them allows us to visualise both metabolic activity and anatomical structure.

2. The relationship between the uncertainty in spatial localisation (Δx) along the line of response and the time difference (Δt) between the detection of the two photons can be described by:

$$\Delta x = \frac{c \cdot \Delta t}{2}$$

where c is the speed of light. If Δt is low enough that Δx is either less than or equal to the detector spatial resolution (around 4-5 mm) then the location of emission is precise enough that image reconstruction is not needed [6]. Using this information we can see that we would need the time difference to be lower than or equal to $\frac{2a}{c}$ where a is the detector spatial resolution. This is incredibly low and close to 0 (we are dividing by the high number of the speed of light) if we use 4-5mm so we need the time difference in our TOF-PEF to be extremely highly accurate.

3. As mentioned previously iterative methods are preferred for PET reconstruction because of the photon attenuation and scatter that you do not get in CT reconstruction. OSEM in particular is the standard because it is faster per iteration and provides a higher quality image in less iterations compared to other iterative methods like MLEM for example (see section 2.5).

   OSEM uses a multiplicative update rule based on expectation-maximisation and uses ordered subsets which allows it to converge faster than gradient descent but is more computationally expensive and may not converge to the true MLE. OSEM also works by iteratively maximising the likelihood whereas gradient descent works by minimising the loss function in the direction of the steepest descent.

4. In PET-MR, you need to process MR data to create an attenuation map for PET, which isn't necessary in PET-CT where CT directly provides this information

3 MRI Image Denoising

3.1 Overview

We need to analyse 3D MRI k-space (frequency space) data from a knee scan acquired using 6 coil receivers. The data is stored in Fourier space and contains artifacts commonly found in MRI imaging, requiring visualization, noise identification, and denoising techniques to improve image quality.

3.2 Visualisation

We load the k-space data and observe that the 1st dimension, corresponding to the coil receivers, has a size of 6, indicating 6 separate receiver coils used in the MRI acquisition. We create 1 image of the magnitude of the K-space for each of these 6 coils, while using log to scale them for enhanced visibility (see the accompanying for these grayscale images). Using the inverse Fourier transform we transpose the raw K-space data into Image space and show a magnitude and phase image from the first coil, shown below:

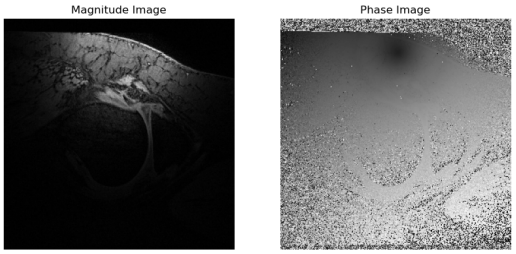

Figure 11: Magnitude and Phase image of the first coil

We take this further in the accompanying code by

displaying the magnitude images for the rest of the coils as well. We finish by combining the data from all of the coils into a single image (in image space), which is shown below:

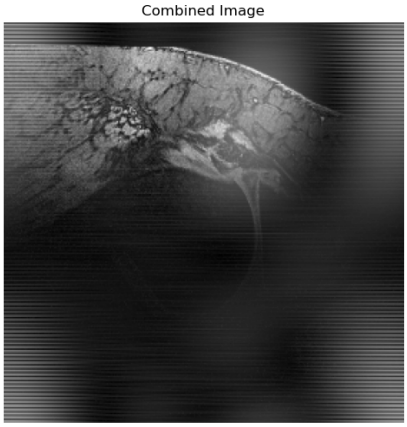

Figure 12: Combined image from all the coils in image space

See the accompanying code for how we did this.

With the combined image there are a lot of lines through the image which are ringing artefacts which could be due to under-sampling, phase inconsistencies but also the addition of all the noise from each coil because we are combining them all into one image. Four of the coils in K-space are fully grey which suggests they have very low variation in signal. The 4th coil is capturing one frequency band and the 6th coil is only receiving a single frequency component. This could suggest that the 4th coil contributed most to the ringing in the combined image as it's vertical k-space line directly translates into horizontal streaks in image space. There are also limitations to the method we used to combine the images and we could have used a more advanced method such as SENSE which uses phase and sensitivity information to construct better images [7]. The last observation is that the magnitude images degrade in quality as you go through the coils until the 6th one is just a blurry image which shows that the later coils contain a lot of noise.

3.3 Denoising

We used three denoising methods on the image space data which were Gaussian, bilateral and wavelet. Firstly we observe that these denoising methods worked better for the coils that had greyscale K-space images. This was observed through viewing that there was not much, if any difference after the denoising was used on coil 4 and 6. For the rest of the coils Gaussian filtering makes the images less noisy but are blurry. Bilateral Filtering preserved the edges, producing a sharper result than Gaussian while maintaining the visual structure. We can use these two fast and reliable methods as a benchmark for the later questions but also for the 3rd denoising method of using wavelets. Wavelets seems to be the most effective method for MRI as it balances noise removal with preserving the important image structures. We can take the 1st coil as an example in this report and all the other coils can be seen in the accompanying code.

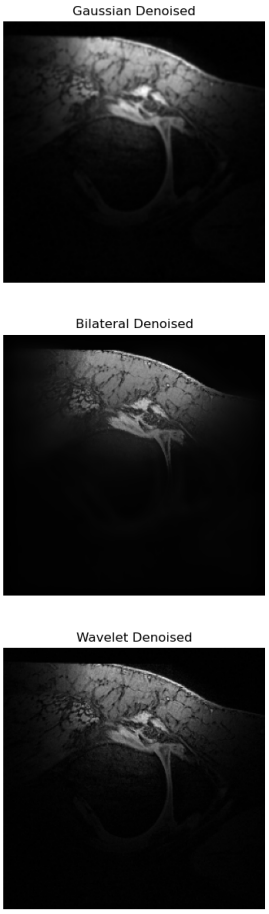

Figure 13: Images after denoising the first coil

We can observe here that the Gaussian has removed noise from the original image space data of this first coil but it does appear very blurry especially at the edges however this method is fast and easy to use however for MRI images it is far from desirable to

implement. Bilateral Filtering we can see has improved on this and gives a sharper less blurry image maintaining clarity especially at boundaries. The Bilateral filtering is more desirable than Gaussian because it maintains the anatomical details and is far less blurry however wavelets as we can see here improves this further. The level of detail in the wavelet image is even clearer to see with less visible regions and boundaries being sharper and more distinguished, as it perseveres the anatomical features. The wavelets method is more desirable for MRI denoising because it gives the better image as we have observed however it is also a much more complex method to implement and to understand.

3.4 Low-pass Butterworth Filter

We apply a low-pass Butterworth filter (within k-space) on the original noisy k-space data and display below the first coils magnitude and phase images with this applied:

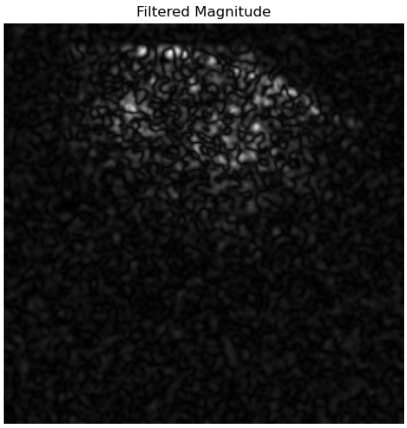

Figure 14: Magnitude image with low-pass Butterworth filter applied to the original noisy k-space data

The Butterworth filter removes high-frequency components, which means fine details and sharp edges are reduced. The image is extremely blurry which is because low-pass filtering in k-space removes high-frequency details, which directly leads to blurring in the image domain.

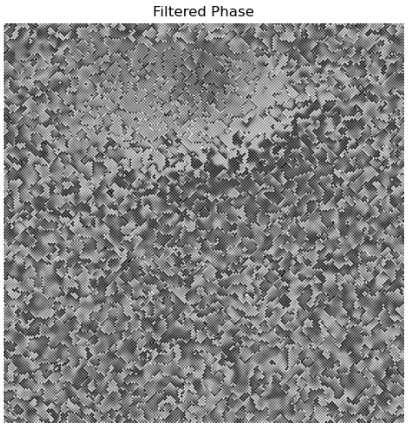

Figure 15: Phase image with low-pass Butterworth filter applied to the original noisy k-space data

This image is also blurry and a poor construction compared to the other methods used previously. Tuning the parameters (D0 and n) has a better impact on the images by increasing D0 or setting n to 1 we get better results than we did with Gaussian but it still falls short of the wavelets and bilateral performances. High frequencies are removed with the Butterworth filter which causes the blurring. The filter affects the Fourier components non-uniformly, which can introduce artifacts leading to the phase image looking messy. For this MRI task the Butterworth filter (with the predetermined parameters) performs poorly but it can be useful for removing extreme high-frequency noise and if some artefacts originate in k-space.

3.5 Combined Image with Wavelets

We use the wavelet denoising method on the images and recreate a combined image using this to observe whether it improves it or not. Despite the improvements from the wavelet denoising, ringing artifacts remain visible. Noise has been removed and the image is clearer while the anatomical features are maintained especially at boundaries however the ringing is still visible due to it being a systematic artefact from the k-space data which is not what the wavelet denoising method is designed to eradicate. The combined image is shown below:

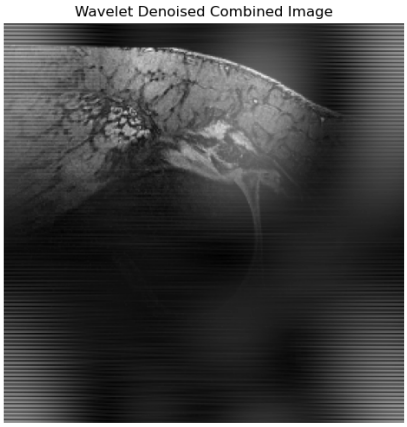

Figure 16: Combined image after wavelet denoising

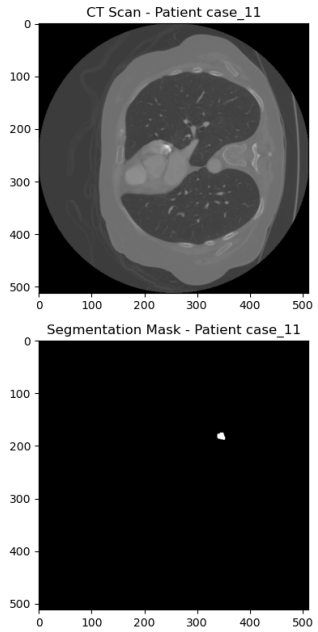

Figure 17: A patients CT scan and segmentation mask

3.6 Further Methods

Our biggest problem seems to be with ringing which occurs if high-frequency k-space data is undersampled so one method would be to increase k-space sampling. In terms of further denoising methods, we could turn to blind image denoising or diffusion denoising for example. For better methods for combining the images we could use the parallel imaging techniques SENSE and GRAPPA or we could weight each coil image based on its Signal-to-Noise Ratio which gives higher importance to coils with better signal quality.

More denoising methods include PCA, NLM and further deep learning methods. Recent works include a way to combine NLM and PCA together [8] which has been shown to even out perform CNNs while having higher PSNR, SSIM, and running speed. Deep neural networks can also be an effective choice of denoising method [9] because they exhibit impressive generalisation ability and show state-of-the-art performance on different datasets.

4 CT image Classification

We have a folder that contains CT scans for a selection of 40 patients of lung cancer.

4.1 Image Segmentation

We open the files and create 1 Numpy array per patient scan and 1 per segmentation mask, we display 1 of the patients images below:

For each patient, we compute the minimum and maximum x, y, and z coordinates where the cancer is present in the segmentation mask. After doing this we create a numpy array with a subvolume of the images (smaller 3D region extracted from the full CT scan). We extract only the region where the tumour is located to make analysis efficient and accurate but we also increase the voxels around where we computed the cancer to ensure none of the cancer lies outside of this as the segmentation mask might only include the main part of the cancer.

Next, we implement a simple threshold-based segmentation function and apply it to the extracted subvolumes. Below, we display a segmented tumour image for a selected patient using our method.

Figure 18: A patients segmented tumour based on our simple thresholding algorithm

Considering we have increased the voxels in the original ground truth ones to over-cover the tumour and the fact we only used a simple algorithm it does not perform too badly. The thresholding approach does not generalise well to different tumour shapes and intensities though, leading to over-segmentation. The ground truth masks are more precise however thresholding approach captures some tumour regions but also includes too much non-tumour tissue (irrelevant segmented regions we might not expect).

To improve on this performance we could use the Otsu algorithm, edge detection, feature extraction or K-means clustering. Ultimately, transitioning to machine learning-based segmentation, such as U-Net or other deep learning models, would provide better generalisation across different tumour shapes and intensities while minimising irrelevant tissue inclusion. A CNN with minimal preprocessing has also been shown to reach the performance close to human observers on brain images [10].

4.2 Image Feature Extraction

We have three histogram-based radiomic features (the mean absolute deviation (MAD), the intensity and the uniformity) which provide insights into the tumour's texture and intensity characteristics. See the accompanying code for implementation details. We compute the value of each feature on each patient's nodule, displaying three patients below:

Patient	MAD	Intensity	Uniformity
0	189.50	27673732	0.0001
1	106.04	57513796	0.0001
10	274.03	6673689	0

The uniformity values are very low for all three, suggesting a widely spread intensity distribution across tumour regions. Patient 0 has high intensity with moderate MAD, indicating significant variation while Patient 10 has the lowest intensity but highest MAD, reflecting the most heterogeneous intensity variation.

MAD is a useful feature for classifying tumours as benign or malignant because it quantifies the variation in voxel intensity within the tumour. Malignant tumours tend to have more heterogeneous intensity distributions due to irregular growth leading to higher MAD values. In contrast, benign tumours are usually more uniform, resulting in lower MAD values. Compared to other features like intensity and uniformity, MAD directly captures variability within the tumour, making it particularly effective for classification.

References

[1] Diego Garcia Pinto, Oliver Diaz... (2018), Breast tomosynthesis reconstruction using TIGRE software tool, Research Gate, Fourteenth International Workshop on Breast Imaging

[2] Michael J. Reinhardt, Nicole Wiethoelter... (2005), PET recognition of pulmonary metastases on PET/CT imaging: impact of attenuation-corrected and non-attenuation-corrected PET images, Springer Nature Link

[3] Urvi Joshi, Pieter G. H. M. Raijmakers... (2007), Attenuation-Corrected vs. Nonattenuation- Corrected 2-Deoxy-2-[F-18]fluoro-D-glucose- Positron Emission Tomography in Oncology, A Systematic Review, Academy of Molecular Imaging

[4] Ashish Kumar Jha , Nilendu C Purandare... (2021), PET reconstruction artifact can be minimized by using sinogram correction and filtered back-projection technique, Indian Journal of Radiology and Imaging

[5] Jens Mincke , Jan Courtyn... (2021), Guide to Plant-PET Imaging Using 11_{CO_2}, Research Gate, Frontiers in Plant Science

[6] Suleman Surti... (2015), Update on time-of-flight PET imaging, National Institute of Health

[7] Thomas Sartoretti... (2018), Common artefacts encountered on images acquired with combined compressed sensing and SENSE, Insights Imaging

[8] Shiao Li... (2024), New non-local mean methods for MRI denoising based on global self-similarity between values, Institute of Medical Technology, Peking University Health Science Center

[9] Zhaoming Kong... (2023), A Comparison of Image Denoising Methods, Cornell University

[10] Ramin Ranjbarzadeh... (2021), Brain tumor segmentation based on deep learning and an attention mechanism using MRI multi-modalities brain images, Nature, Scientific Reports