IMPERIAL

Final report for MEng Project

Pseudo-Labelling with Pairwise Image Registration:
A Unified Framework for Motion Correction in Super-Resolution
Ultrasound Imaging

Elliot Ocanagil-Tunstall

Supervisor(s):

Professor Mengxing Tang

Clara Rodrigo Gonzalez

Submitted in partial fulfilment of the requirements for the award of MEng in Molecular
Bioengineering from Imperial College London

Table of Contents

Abstract

Accurate imaging of microvasculature can improve clinical outcomes. However, conventional ultrasound (US) cannot resolve microvasculature structures at clinical frequencies due to the fundamental diffraction limit. Super-resolution (SR) overcomes this limitation by localising and tracking contrast agents across multiple frames, requiring seconds to minutes of acquisition. This makes SR US techniques prone to motion artifacts. Motion correction determines the achievable resolution limit, but current methods have yet to reach the target accuracy required to image capillaries (< 10um). Image registration can be employed for motion correction, but the optimal parameters for US applications are not well-documented. The low contrast and high variability of US images make registration challenging. However, segmentation techniques could help overcome these limitations. This project aims to improve motion correction accuracy for US using pairwise segmentation and registration. A semi-automatic segmentation method is used to accurately segment regions of interest (ROI), achieving a Dice score > 0.95. These ROI masks and images are simultaneously aligned using diffeomorphic image registration, resulting in a 132% increase in anatomical alignment and a 69.3% decrease in computation time, compared to a best-in-class image registration algorithm. These results align with findings from other studies, though direct comparisons are difficult due to the significant variation in deformation across studies. This work provides a unified framework for motion correction in SR US and underscores the importance of using masks in image registration. This insight could lead to novel region-specific approaches to motion correction in SR US imaging.

1. Introduction

1.1. Motivation

The role of the vascular system is crucial in the pathology of various diseases, ranging from neurodegenerative disorders of the brain, such as Parkinson's disease [1], to cancers across numerous tissue types. Understanding vasculature brings a unique perspective to pathology, for instance the characterisation of blood vessel formation (angiogenesis). Increased pathologic angiogenesis is an early event in tumour development [2], characterised by an irregular microvasculature lacking normal hierarchical branching [3]. Imaging these characteristic vascular formations could significantly enhance early tumour detection and improve prognoses for cancer patients [4]. Angiogenesis is also prominent in cases of arteriosclerosis, which can result in strokes and cardiac arrest [5]. The capacity to visualise microvascular systems offers distinct clinical advantages, facilitating earlier interventions and potentially better treatment outcomes. However, due to the limitations of imaging technologies at microscopic scales, imaging microvasculature such as capillaries, typically measuring less than 10μm in diameter, has presented a significant challenge and now represents the target resolution for future imaging innovations [6].

Ultrasound (US) is an effective technique for vascular imaging and is widely favoured by clinicians for its real-time acquisition capabilities and ease of use. Its non-ionising nature provides an advantage over computed tomography (CT), and its cost efficiency makes it favourable over to MRI. Despite its clinical benefits, US's ability to resolve sub-wavelength structures, such as vascular capillaries, at a clinical frequency is fundamentally limited by the acoustic diffraction limit ($\lambda/2$), approximately 150μm for a 5MHz US wave [6]. The last two decades of research have seen huge progress in overcoming this limitation, with the emergence of an exciting and promising technique: super-resolution (SR) US imaging.

1.2. Super-Resolution Ultrasound Imaging

SR US relies on contrast agents to visualise sub-wavelength vasculature that is otherwise unresolvable. These agents, typically microbubbles (MB), are injected intravenously and circulate through the vasculature. As they pass through, they resonate and scatter the incident ultrasound waves non-linearly, enabling easy identification while enhancing the contrast between blood and surrounding tissues [7]. Ultrasound localisation microscopy (ULM), as visualised in Figure 1, utilises the contrast provided by MBs and is the overarching technique used to create SR images. It leverages the cumulative localisations of numerous discrete MBs across a vast number of frames, which has enabled the reconstruction of a super-resolved composite image with high spatial resolution [8]. However, it can take several minutes for these MBs to traverse the entire vasculature. As a result, tracking localised MBs is highly susceptible to motion disturbances, often leading to an inaccurate mapping of the underlying vasculature. Most of the body's natural motion is rigid due to the diaphragm moving 15mm during breathing [9]. Non-rigid cardiac motion can also have a substantial effect on periphery organs, with the region of the liver adjacent to the heart capable of moving up to 4mm [9] within a cardiac cycle. Predictable large global motion can be compensated for. An internal study used an impedance pneumometer to only image during the respiratory rest period [10]. Other techniques use US image intensity to phase the heart into its distinct cardiac cycles, separating frames with large motion [11]. Small motion, along with unavoidable motion, such as gut peristalsis, coughing and bowel movements [12] are more difficult to correct. To solve this, motion correction algorithms are implemented prior to MB detection. Motion currently determines the achievable resolution limit for SR US imaging [12], and hence to reach the desired objective of imaging capillaries, we require motion correction accurate to < 10μm. This level of accuracy has yet to be achieved reliably [6].
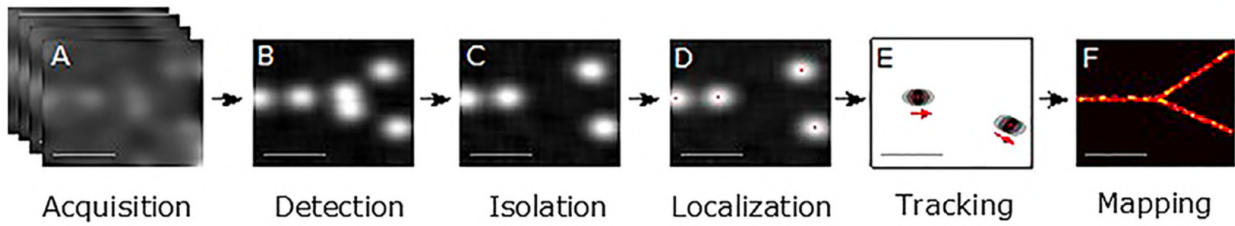
Figure 1: Steps in ULM: (a) Acquisition of US data using contrast agents. (b) Detection of MBs contrast agents. (c) Isolation of individual signals by rejecting overlapping. (d) Localisation of the MBs using SVD. (e) Tracking MBs over time to determine velocity profile. (f) Mapping of the accumulated localisations to produce a super-resolved composite image with high spatial resolution. Extracted from [6].

Tissue segmentation is another important post-processing step. For heart and gut acquisitions, identifying a region of interest (ROI) enables computationally expensive microbubble (MB) localisation processes, such as singular value decomposition (SVD), to be applied with regional necessity. Segmentation is essential in other US methods, such as myocardial perfusion quantification (MPQ), which is instrumental in assessing myocardial blood flow and monitoring cardiovascular disease (CVD) [13]. Myocardial contrast echocardiography (MCE) is used to visualise MBs [14], where the contrast intensity serves as a reliable indicator of myocardial blood flow. The accuracy of this technique depends critically on the segmentation of the myocardium [11]. Manual annotation is often required to achieve high accuracy which is a laborious and costly process for experienced clinicians. There is a need for accurate segmentation and motion correction for US imaging. The augmentation of these methods into the SR US processing pipeline is visualised in Figure 2.
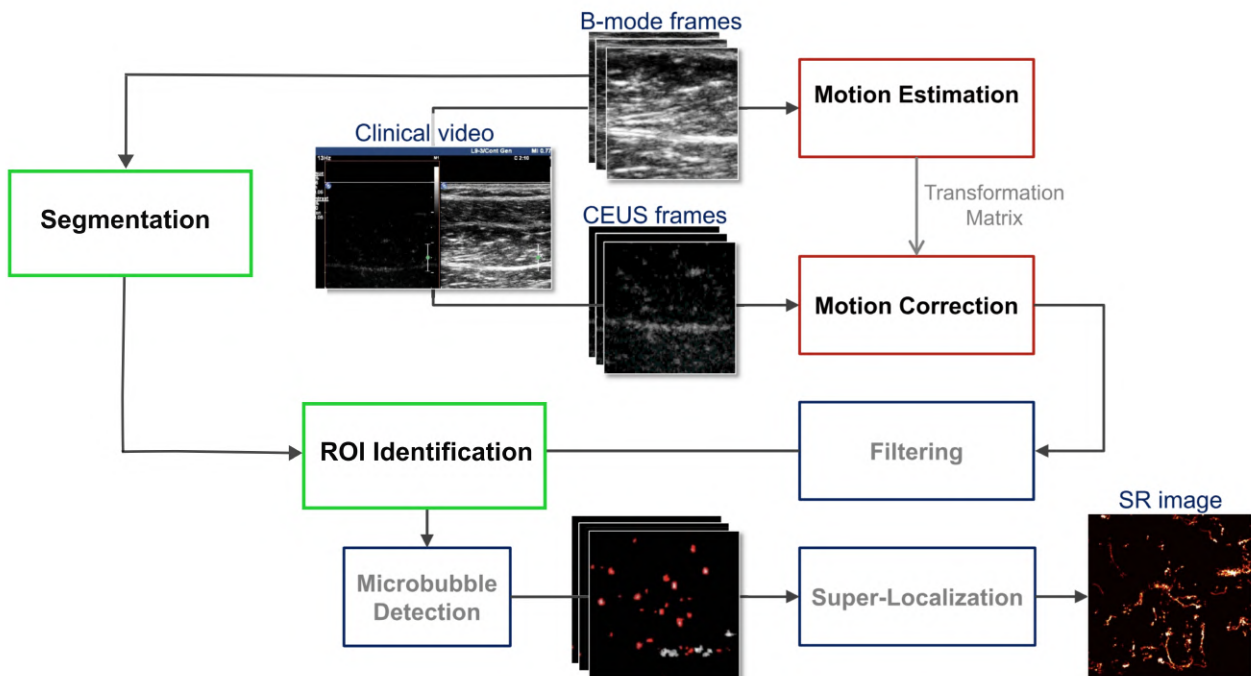

Figure 2: Processing pipeline for SR US imaging. Motion correction processes (red) and segmentation processes (green) are augmented into the SR processing chain. Modified from [12].

1.3.   Diffeomorphic Image Registration

Image registration can be used for both motion correction and image segmentation. This technique uses transforms to map points from a moving image to a reference image (fixed) to align their distinct features with a common coordinate system. Transforms can be classified as rigid (translation and rotation only), affine (rigid + scaling and shearing), or non-rigid (capable of local complex deformations).

Diffeomorphic image registration is a framework for non-rigid image registration that retains image topology by resampling one deformation field by another [15]. Its ability to make small but also large diffeomorphisms whilst retaining the individual spatial relationship of an image make it advantageous for medical image registration [16]. A diffeomorphism is a smooth, continuous, globally bijective transform whose derivatives are invertible [15]. This leads to well-defined, realistic output images that preserve their volume. Being invertible is another key advantage. Inverse transforms support pseudo labelling and the validation of image registration through the comparison of estimated segmentation masks with a ground truth. This is the foundation for a technique called semi-automatic image segmentation which will be explored further in Section 1.5.

State of the art methods of diffeomorphic image registration are symmetric in respect to their image inputs [17]. These exploit the availability to generate an inverse transform by breaking the transform in two. Forward and reverse transformations are optimised simultaneously by minimising a cost function [18]. The cost function balances image similarity with deformation smoothness by incorporating a regularisation penalty term, ensuring that neighbouring transformation points vary smoothly across the spatial domain. One implementation of symmetric diffeomorphic image registration is Dipy [19]. An unpublished internal study found that Dipy was among the most effective open-source algorithms for motion correction in US [20]. Results from this study are shown in Figure 3.
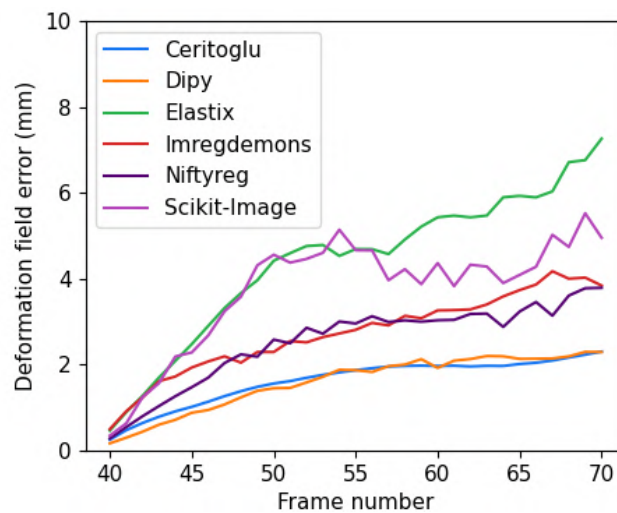

Figure 3: An evaluation of the best image registration algorithms for motion correction in US. Sourced with permission from [20].

1.4.   Motion Correction in Ultrasound

Motion correction is already common practice in many forms of medical imaging and is typically done using non rigid image registration; however, this method alone is not sufficient to reach the accuracy required to image capillaries by SR US. Recent work by S.Harput and M.Tang was able to improve upon traditional techniques using a two-stage image registration approach [12]. Affine transformations were first used for global motion estimation, followed by non-rigid B-spline-based free-form deformation (FFD) image registration for local tissue deformations. This two-stage method was shown to achieve a mean average error that was 20% smaller than solely non-rigid methods [12]. Although this work saw a great improvement, the problem with using FFD method, such as B-spline, is that they do not ensure topology preservation, unlike diffeomorphic methods [21]. Some modern techniques use ML for motion correction. A group at the University of Texas modified a two-stage motion correction method used for MRI to train a fully convolutional neural network (FCNN) [22]. The model was able to produce results of similar accuracy to other two-stage algorithms in 94.8% less time. Although promising, these techniques are limited by their reliance on training data, where a ground truth must be provided by image registration algorithms or realistic simulated data with known deformation fields [23]. Generating accurate ground truth measures can be challenging for US images.

US images are low contrast and are subject to a characteristic and prominent speckle pattern, resulting in higher image variance than other modalities such as CT and MRI. This makes US images significantly more challenging to register accurately. The characteristic speckle pattern hinders intensity-based metrics which often lead to anatomical dissimilarities post registration [24]. Therefore, carefully selected smoothing and optimisation parameters become crucial for minimising the effects of speckle. However, the optimal parameters for US are often unknown due to limited relevant literature, and the predominant focus of open-source image registration methods on MRI. Within the context of SR US, we are more concerned with the anatomical spatial similarity of the images rather than the correspondence of pixel intensities. Intensity-based measures are thus unsuitable metrics. Their susceptibility to image artefacts, such as speckle, and their requirement for fine-tuning of parameters lead to inaccurate registrations that are unrepresentative of the underlying anatomical structure. Binary masks, generated via image segmentation, can be used to align regional features [24] and are less susceptible to image variation and parameter optimisation than images. They could be used to overcome current limitations of motion correction in SR US.

1.5.   Image segmentation

ML-based techniques, such as FCNNs, have been successful in US image segmentation. Mishra et al. created an FCNN where fine levels were used to capture object boundaries and coarse layers used to distinguish object regions from the background [25]. It was tested on blood vessel segmentation in the liver and outperformed a U-Net, the current leading approach, achieving an average Dice score (DS) of 0.83. This score indicates a high level of accuracy in matching the segmented blood vessels with the ground truth.

There are numerous other automatic segmentation methods that are effective for specific types of images. However, in the case of ultrasound imaging, traditional methods such as thresholding and boundary contour fitting, often fall short due to inherent noise and speckle patterns. [26]. Of traditional segmentation techniques, K-means clustering performs best, achieving an accuracy similar to FCNNs for simple segmentation tasks [26]. Semi-automatic segmentation (SAS) methods are particularly intriguing. After a clinical US acquisition, it is common for an experienced clinician to segment a single frame manually. By registering all frames to the annotated frame and applying the corresponding inverse transform to the segmentation mask, we can propagate the segmentation across the entire dataset. This process is also known

as pseudo-labelling. S.Sze and M.Tang used diffeomorphic demons image registration with mean squared error (MSE) optimisation to show that SAS of the myocardium performs in line with other state-of-the-art segmentation approaches, achieving a DS of 0.82 [11]. Registration can also be applied to improve other image segmentation methods. Prevost and Ardon were able to improve kidney segmentation in contrast-enhanced ultrasound (CEUS) by employing a joint-segmentation and registration approach [27]. They integrated registered 3D ultrasound data with CEUS prior to segmentation with a random forest algorithm. This method significantly outperformed existing state-of-the-art techniques, reducing the kidney volume error by half [27]. Registration is significant in multimodal imaging applications, serving as an effective pre-processing method for aligning images from different modalities [28].

The benefits of a joint registration-segmentation approach are clear; image registration improves segmentation techniques. But could this relationship be symbiotic? Until now, there has been little research into the utilisation of segmentation masks for US motion correction.

1.6.   Aims and Objectives

The objective of this project was to develop a unified framework for image registration and segmentation in US imaging, with an overall goal of improving motion correction in US images. The project had three key objectives:

1. To develop a well-defined and robust SAS method for ROI segmentation in US images, using symmetric diffeomorphic image registration.

2. To develop a novel approach to image registration that optimises the alignment of the ROI masks and images simultaneously, using symmetric diffeomorphic image registration.

3. To show that a novel mask-derived approach outperforms current leading symmetric diffeomorphic image registration algorithm in measures of anatomical alignment.

2.   Methods

2.1.   Data Generation

Two simulated US datasets were generated internally by the research group. For each simulated image, a set of tissue sample locations or the resultant deformation field is known. These provide the ground truth for the deformation and allows us to verify the accuracy of motion correction techniques.

Cardiac Images

A simulated tissue phantom was designed with a density of 10 scatterers per resolution cell for each simulation ensuring the generation of a fully developed speckle pattern [29]. Realistic motion patterns were generated through spatial and temporal interpolation of points cross 30 consecutive MRI images of the heart, from the STRAUS dataset [30]. The resulting non-rigid movement of points represents a model of the physical deformation of the heart throughout the cardiac cycle. The deformation at a single point in time is displayed in Figure 4 (left). Simulated tissue deformation was generated by mapping the scattering points to ground truths from the MRI deformation model. Every scatter is transformed by 3D non rigid motion.
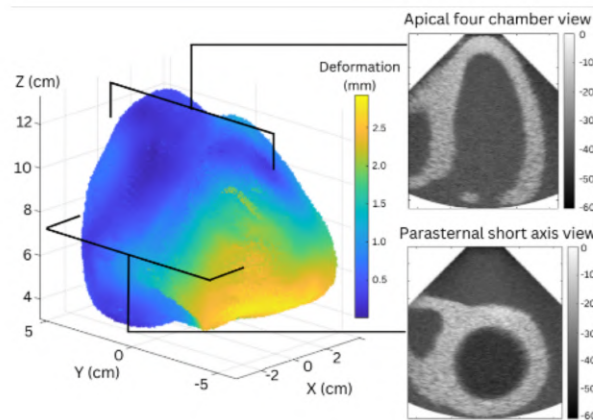

Figure 4: A realistic deformation model of the heart visualised at a single point in time (left) and the resultant simulated images representing the deformation model at the same point in time (right). Sourced with permission from [20].

Cardiac RF images were generated using BUFF [31], a wrapper for the ultrasound simulation platform Field II [32]. BUFF allows for the additional simulations of microbubbles within US images. 165 frames were generated for a full cardiac cycle, resulting in a 220Hz frame rate. Beamforming was used to create the final B-Mode images. 5 transducer angles were used and an additional inter-angle rigid motion term (-0.4mm - 0.6mm) added for clinical realism. Two of these angles can be seen in Figure 4 (right). A parasternal short axis view was chosen for the analysis as this view had maximal deformation.

Carotid Images

Clinical B-Mode images of the carotid artery were used as the basis for the simulated images. These B-Mode images were first segmented using thresholding, with the intensity modelled by a 2D probability density function (PDF) to represent the likelihood of scatterer locations within the imaging field. Subsequently, realistic scatterer locations were generated through Monte Carlo sampling of the PDF. A non-rigid movement model, shown below, was used to create non-rigid deformations of the scatterers, with deformation parameters randomised between images. This model excludes y-plane deformation to negate out-of-plane motion.

Displacement field, a combination of polynomials:

$$d_i(x, y, z) = a_x P^{(N_x)} x + 0y + a_z P^{(N_z)} z$$

Resultant displacement, a weighted sum of the displacement fields:

$$d_r(x, y, z) = b_x d_x(x, y, z)\hat{x} + 0\hat{y} + b_z d_z(x, y, z)\hat{z}$$

where $i \in \{x, y, z\}$; a_i and b_i are random constants; $P^{(N_i)}$ is a polynomial of random order.

For each clinical image, 24 frames were simulated using the Field II and beamforming techniques. The simulations used a single plane wave, a 128-element transducer array, no apodisation, and a central frequency of 5MHz.

2.2.   Automatic Image Segmentation

Manual image annotation was not achievable within the time scale of this project and thus, automatic image segmentation was required. A segmentation toolkit was developed to provide automatic image segmentation for any US image, generating a list of tissue and background objects sorted by size. The toolkit includes various traditional segmentation methods and post processing techniques derived from relevant literature. This allowed for experimentation to identify the most appropriate tools for different datasets. The tools included are shown in Table 1.

Segmentation		**Post Processing**	
Method	**Implementation**	**Method**	**Implementation**
Otsu Thresholding	Scikit-image [33]	Gaussian Filtering	OpenCV [34]
K-means Clustering	ANTsPy [35]	Binary Dilation and Erosion	OpenCV
Connective Components Labelling	Scikit-image	Binary Hole Filling	SciPy [36]
Regional Property Measurements	Scikit-image	Contour Smoothing	OpenCV

Table 1: The image segmentation toolkit. Development tools for image segmentation (left) and development tools for post processing (right).

After segmentation, individual tissue and background regions were separated using connected components labelling with a connectivity of 2. This was followed by regional property measurements for each labelled object and sorting of the binary regions by size. Every 5th mask was validated by visual inspection and deemed accurate to a similar standard as manual annotation. This sample size was deemed large enough to represent the population.

Cardiac Images

60 cardiac frames that were automatically segmented into tissue or chamber using tools selected from the toolbox. K-means clustering segmentation was found to be the most effective segmentation method. The smoothing parameter (mrf) was chosen to be 0.15 through experimentation. A morphological binary closing operation, a binary dilation followed by a binary erosion, using a square kernel of size of 5x5 was used to fill small gaps in the masks. Example segmentations are displayed in Figure 5 with tissue segmentation in green and chamber in blue for ventricular diastole (left) and ventricular systole (right).
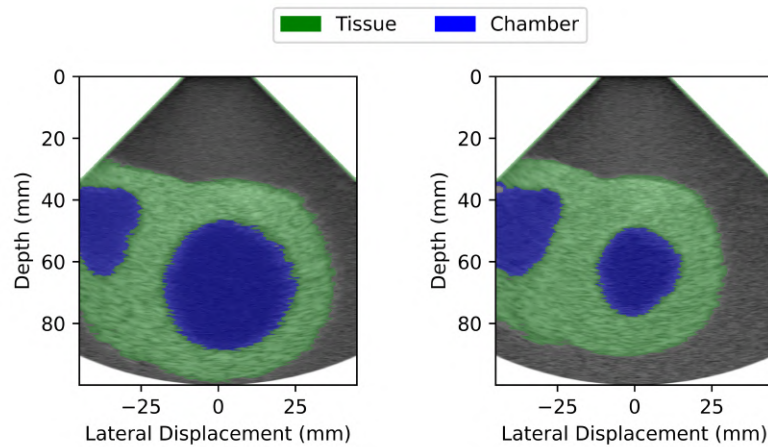

Figure 5: Automatic segmentation of the cardiac images. Frame 1 - ventricular diastole (left) and frame 50 - systole (right). Legend presents green overlay as tissue and blue overlay as chamber.

Carotid Images

25 frames were segmented into tissue and background automatically. Figure 6 shows frame 1 (left) and 25 (right). Otsu-thresholding was used for segmentation followed by various post-processing methods. First, 3 iterations of binary dilation using a cross shaped kernel of size 3x3 was used to connect regions of proximity. This was followed by gaussian smoothing using a 5x5 kernel with a smoothing coefficient (σ) of 1. The aim of this was to smooth rough edges created by the dilation process. Finally, a binary fill operation was used to fill holes within the image.
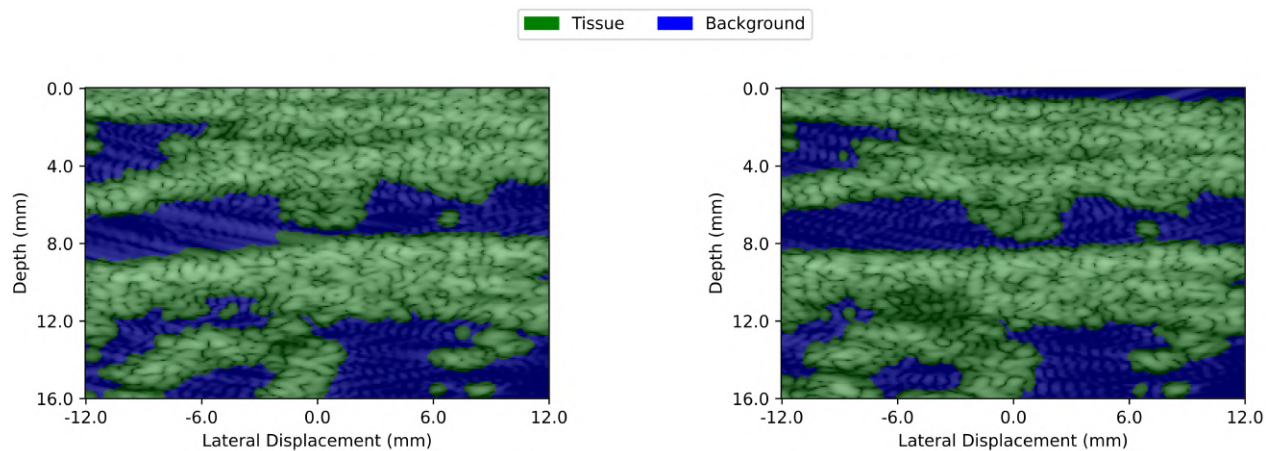

Figure 6: Automatic segmentation of the carotid images. Frame 1 (left) and frame 25 (right). Legend presents green overlay as tissue and blue overlay as background.
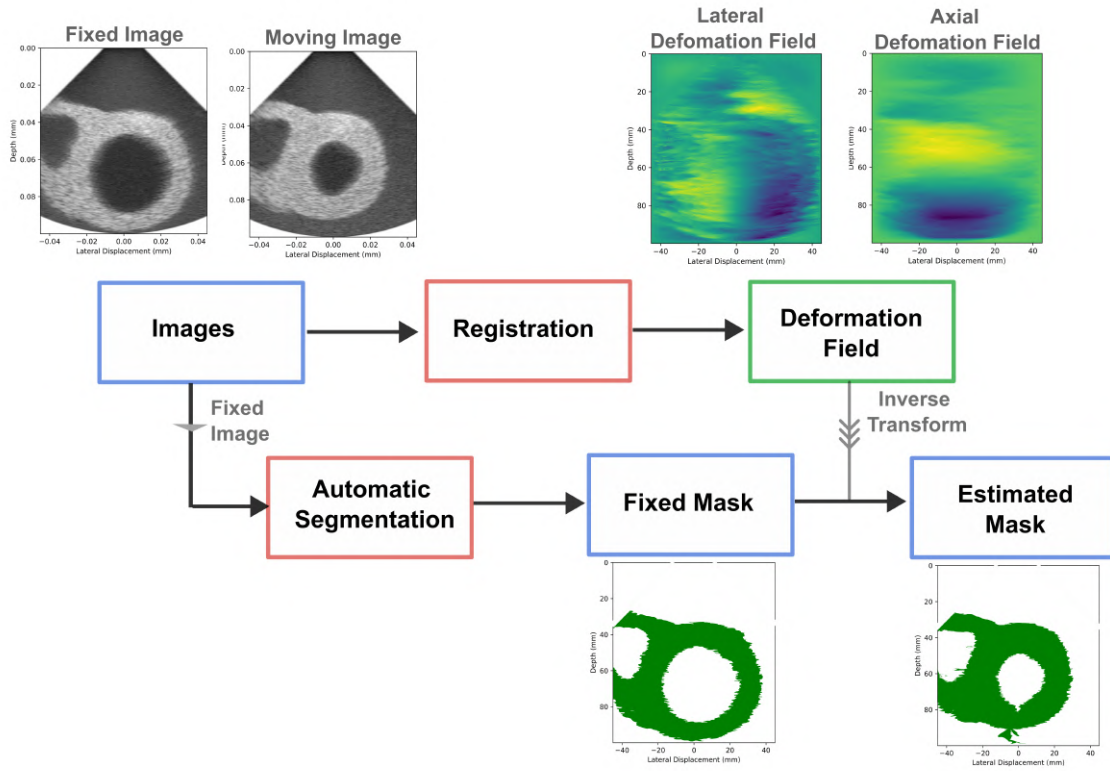
2.3.   Semi-Automatic Image Segmentation

Figure 7: The SAS methodology where B-Mode images are processed into binary masks. Images and mask objects are shown in blue, processes in red and intermediary stages in green.

The SAS methodology can be visualised in Figure 7. B-Mode images were first registered using Dipy. Dipy was chosen due to its established performance for motion correction. It is written in Python and is open source, allowing for ease of customisation and implementation with existing codebase. Bayesian optimisation was used to find the optimal registration parameters for a single frame to the reference. These parameters were used for all frames in the dataset and are given in Table 2.

	Datatype	
Parameter	**Cardiac**	**Carotid**
Optimisation Threshold (minimum energy limit for optimisation completion)	1.04×10^{-3}	5.81×10^{-4}
Similarity Metric (selection of optimisation metric)	SSD	SSD
Number of Iterations (limit on number of optimisation cycles)	308	338
Number of Pyramidal Layers (consecutive layers in a gaussian scale space)	7	2
Smoothing Coefficient (σ) (smoothing of the deformation field)	0.181	0.368

Table 2: Registration parameters for Dipy obtained via Bayesian optimisation. Parameter descriptions given in blue.

A reference frame, selected from the start of the deformation period, was automatically segmented to create a reference mask. All frames in the dataset were registered to the reference frame, using Dipy. The respective inverse transformations were applied to the reference mask for each registration. To evaluate the accuracy of this segmentation method, the estimated masks were compared against the ground truth masks, generated by automatic segmentation. The following accuracy measures were used:

Dice Similarity Coefficient (DS)

The DS represents the spatial overlap between 2 image sets [37]. It is often used to provide an accuracy for image segmentation.

$$DS(A, B) = \frac{2 \times |A \cap B|}{|A| + |B|}$$

where A and B represent sets of binary images

Hausdorff Distance (HD)

The HD is defined as the maximum distance of a set to the nearest point in another set [38]. It is a useful metric as it highlights large discrepancies in mask shape. Using the image mapping, the distance in was calculated in displacement rather than pixels.

$$HS(A, B) = \max_{a \in A} \left\{ \min_{b \in B} \{ d(a, b) \} \right\}$$

where a is a point in set A, b is a point in set B and d is the distance measure.

2.4.   Co-Registration Algorithm

Dipy was customised to develop a bespoke algorithm for pairwise image registration, referred to as co-registration, as shown in Figure 8. This algorithm accepts images and corresponding masks, creating a multi-resolution scale space for each input using Gaussian filters. At each resolution level, optimisation continues until the energy threshold is met or the iterations limit is reached.

Within optimisation, displayed within the red box in Figure 8, both images and masks undergo identical transformations using the current deformation field. Subsequently, mask and image alignment is independently evaluated using the SSD method, described by Algorithm 1. This approach, based on Vercauteren's efficient non-parametric diffeomorphic image registration method [39] and Thirion's demons algorithm [40], forms the basis of Dipy's optimisation methodology. The final cost function, shown in blue in Algorithm 2, is calculated by combining the SSD output for both masks and images based on parameter weights. By adjusting the weights, the co-registration optimisation can be tuned towards image (α) or mask (β) alignment. The final step in each metric iteration is the gaussian regularisation of the resulting deformation field to ensure smoothness, stability, and realistic deformation in the final output. The registration parameters were the same as those used for the Dipy for SAS, shown in Table 2.
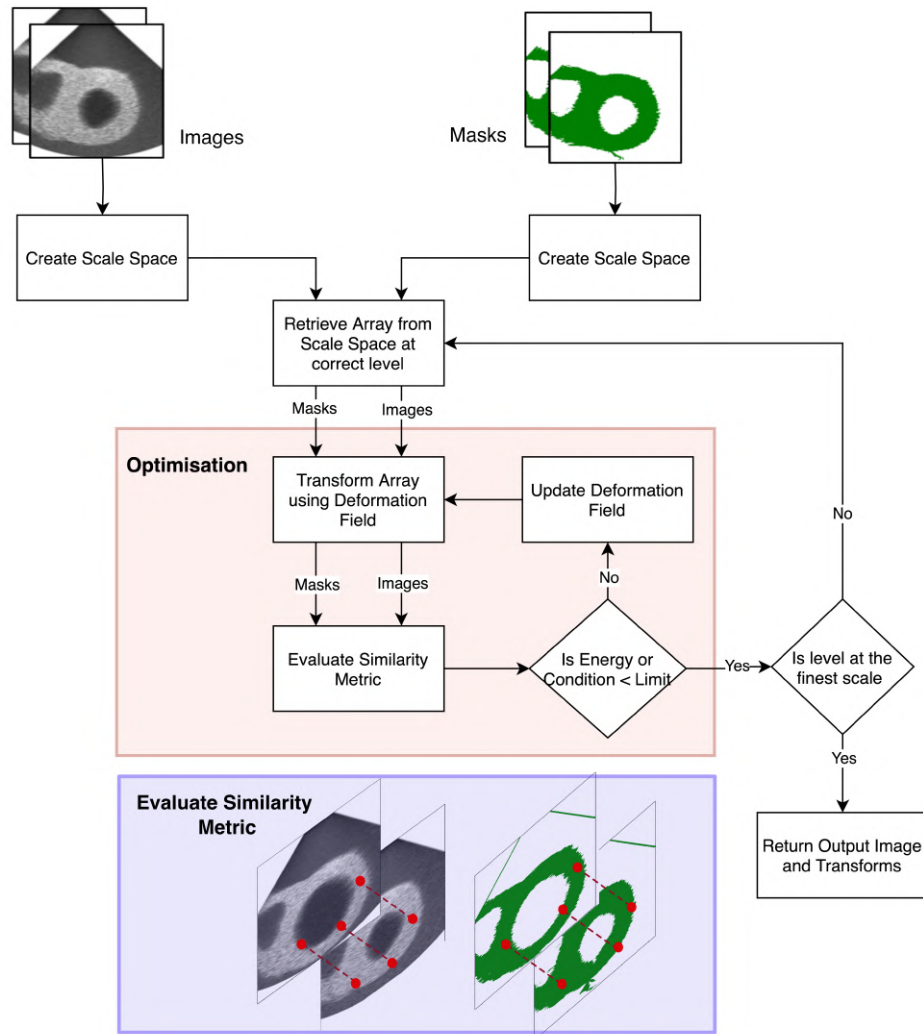

Figure 8: Co-registration algorithm flow diagram. Flow starts from images (top) and masks (top) until the output images and transforms are returned (bottom right). Red box highlights optimisation subroutine. Blue box visualises the simultaneous evaluation of masks and images within the similarity metric.

Algorithm 1 SSD Demons Optimisation Function

1: **Inputs:** $fixed, moving, \sigma_{sq\_x}$
2:
3: Initialize energy: $energy \leftarrow 0$
4: Define delta: $\delta_{field} \leftarrow fixed - moving$
5: Define gradient: $gradient_{moving} \leftarrow gradient(moving)$
6:
7: **for** $i \leftarrow 1$ **to** $nrows$ **do**
8:     **for** $j \leftarrow 1$ **to** $ncols$ **do**
9:         $\delta \leftarrow \delta_{field}[i, j]$
10:         Accumulate energy: $energy \leftarrow energy + \delta^2$
11:         $norm2 \leftarrow gradient_{moving}[i, j, 0]^2 + gradient_{moving}[i, j, 1]^2$
12:         Compute denominator: $den \leftarrow \frac{\delta^2}{\sigma_{sq\_x}} + norm2$
13:         **if** $den < 1 \times 10^{-9}$ **then**
14:             Set output vector to zero: $out[i, j, 0] \leftarrow 0, out[i, j, 1] \leftarrow 0$
15:         **else**
16:             Calculate output vector:
17:             $out[i, j, 0] \leftarrow \frac{\delta \cdot gradient_{moving}[i,j,0]}{den}$
18:             $out[i, j, 1] \leftarrow \frac{\delta \cdot gradient_{moving}[i,j,1]}{den}$
19:         **end if**
20:     **end for**
21: **end for**
22:
23: **Return:** $energy, out$

Algorithm 1: Pseudo code for SSD demons optimisation method implemented by Dipy [19].

Algorithm 2 Co-Registration Metric

1: $energy_{mask}, out_{mask} \leftarrow SSD_{demons}(fixed_{mask}, moving_{mask}, \sigma_{sq\_x})$
2: $energy_{image}, out_{image} \leftarrow SSD_{demons}(fixed_{image}, moving_{image}, \sigma_{sq\_x})$
3:
4: Calculate total energy: $energy \leftarrow \alpha \cdot energy_{image} + \beta \cdot energy_{mask}$
5: Calculate total deformation field: $out \leftarrow \alpha \cdot out_{image} + \beta \cdot out_{mask}$
6:
7: Regularisation: $out \leftarrow gaussianFilter(out, \sigma_{sq\_x})$

Algorithm 2: Pseudo code for the co-registration metric. Final cost function is displayed in blue.

To measure the accuracy of the registration, both absolute error and intensity-based similarity measures were used. For intensity measures the output image from the registration was compared against the fixed image. Three metrics that are common within image registration were used:

$$SSD(A, B) = \sum (A - B)^2$$

Mutual Information (MI), a measure of statistical dependence [41]:

$$MI(A; B) = H(A) - H(A|B)$$

where $H(i)$ is the entropy.

Normalised cross correlation (NCC), which evaluates the correlation between pixel intensity values of two images [42]:

$$NCC(A, B) = \frac{\sum_{i=1}^{N}(A_i - \bar{A})(B_i - \bar{B})}{\sqrt{\sum_{i=1}^{N}(A_i - \bar{A})^2 \sum_{i=1}^{N}(B_i - \bar{B})^2}}$$

The absolute error is a term used in literature for the absolute distance error from the ground truth [12]. For the cardiac dataset, the ground truth displacement for each scatter is known. First, the forward transform generated by the registration is applied to the scatters and then the field of view trimmed by deleting all scatter location outside of the sector field, as visualised in Figure 9. Finally, the distance error between the final scatter locations and known ground truth displacement was calculated. For carotid data, the ground truth is the deformation field. The absolute error is simply the resultant difference between the vector field for every interpolated position. The mean absolute error and its standard deviation was found for each image.
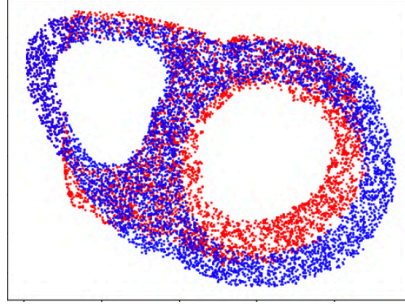

Figure 9: Visualisation of transformed moving image scatters after registration (blue) with scatters of fixed image (red). Source with permission from [20].

2.5.   Pipeline Overview

Figure 10 below shows the overall pipeline for the unified framework. The process starts with the images, highlighted in red, which are used to generate our true masks by automatic segmentation. The true mask of a single reference frame is used to create our estimated masks by SAS, using Dipy. True and estimate masks are then compared using the DS and HD to evaluate segmentation performance. The estimated masks and images are then co-registered using our custom algorithm, producing the final output image.
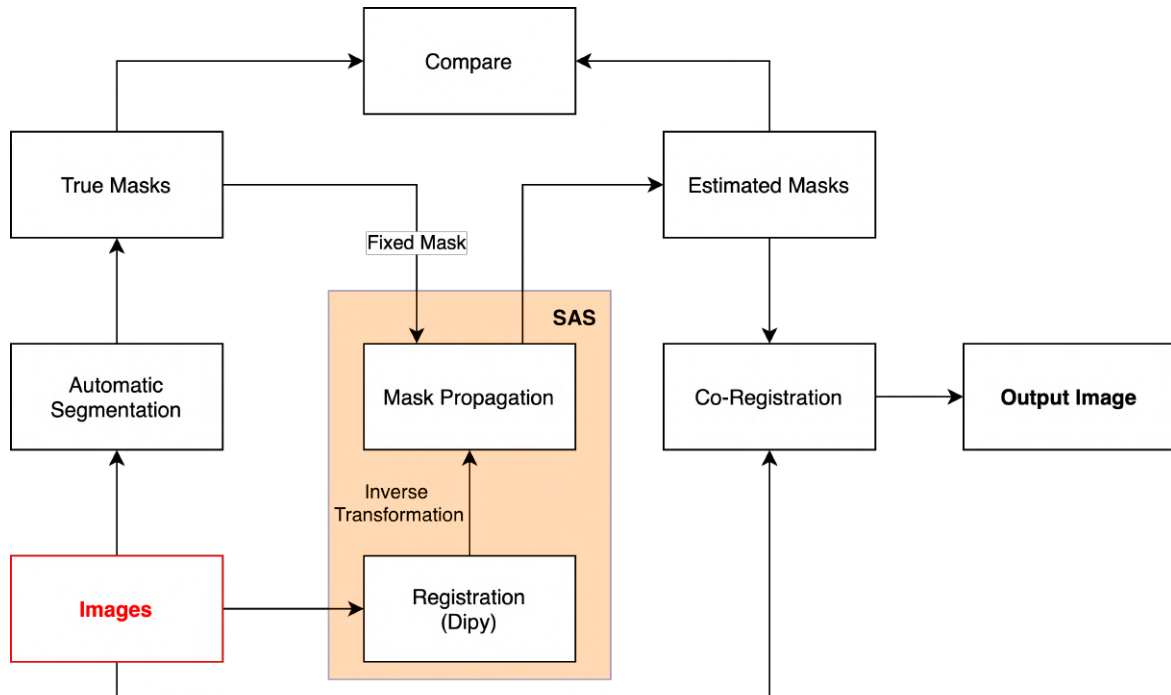

Figure 10: Overview of the pipeline for SAS and co-registration. The process starts with the fixed and moving images (red) and finishes with the output image (right). SAS subprocess is contained within orange box.

3.   Results

3.1.   Semi-Automatic Segmentation

Estimated masks from SAS are evaluated against the ground truth segmentations in Figure 11b. DS (top) and HD (bottom) are plotted against the relative frame number to visualise SAS's accuracy across the entire dataset, which spans a representative deformation for an acquisition period. The relative frame number indicates the position of a frame following the selected reference frame (frame 0). The cardiac cycle is split into its 2 main phases, systole and diastole, to decrease the overall motion and replicate techniques currently used within US post processing [11]. The reference frame was selected at the initialisation of each deformation period. Figure 11a shows the true mask in green, the estimated mask in dark orange and areas of overlap in yellow. These masks correspond to frames sampled from the graphical evaluation in Figure 11b. Top shows the cardiac systole phase, middle is the cardiac diastole phase and bottom is the carotid.

Upon analysis of the segmentation frames in Figure 11a, it was observed that all segmentations for the diastole phase are incomplete. The selected reference frame led to a poor automatic segmentation, limiting the SAS method in its ability to create accurate estimated masks. For this reason, it was decided to omit the diastole data from the remaining analysis.

In Figure 11b, the systole cardiac phase and carotid segmentation show similarities in their DS, with both showing a linear decrease across the dataset with an average DS > 0.95. This is seen in the systole segmentations in Figure 11a, where frame 18 (large deformation) shows the largest proportion of green (false positive) and dark orange (false negative) segmentation regions. In terms of HD, the carotid dataset showed high anatomical accuracy, achieving a maximum displacement of 1mm, while cardiac systole showed a greater displacement of 4.12mm. Overall, the HD increased smoothly with frame number in line with the increasing deformation.
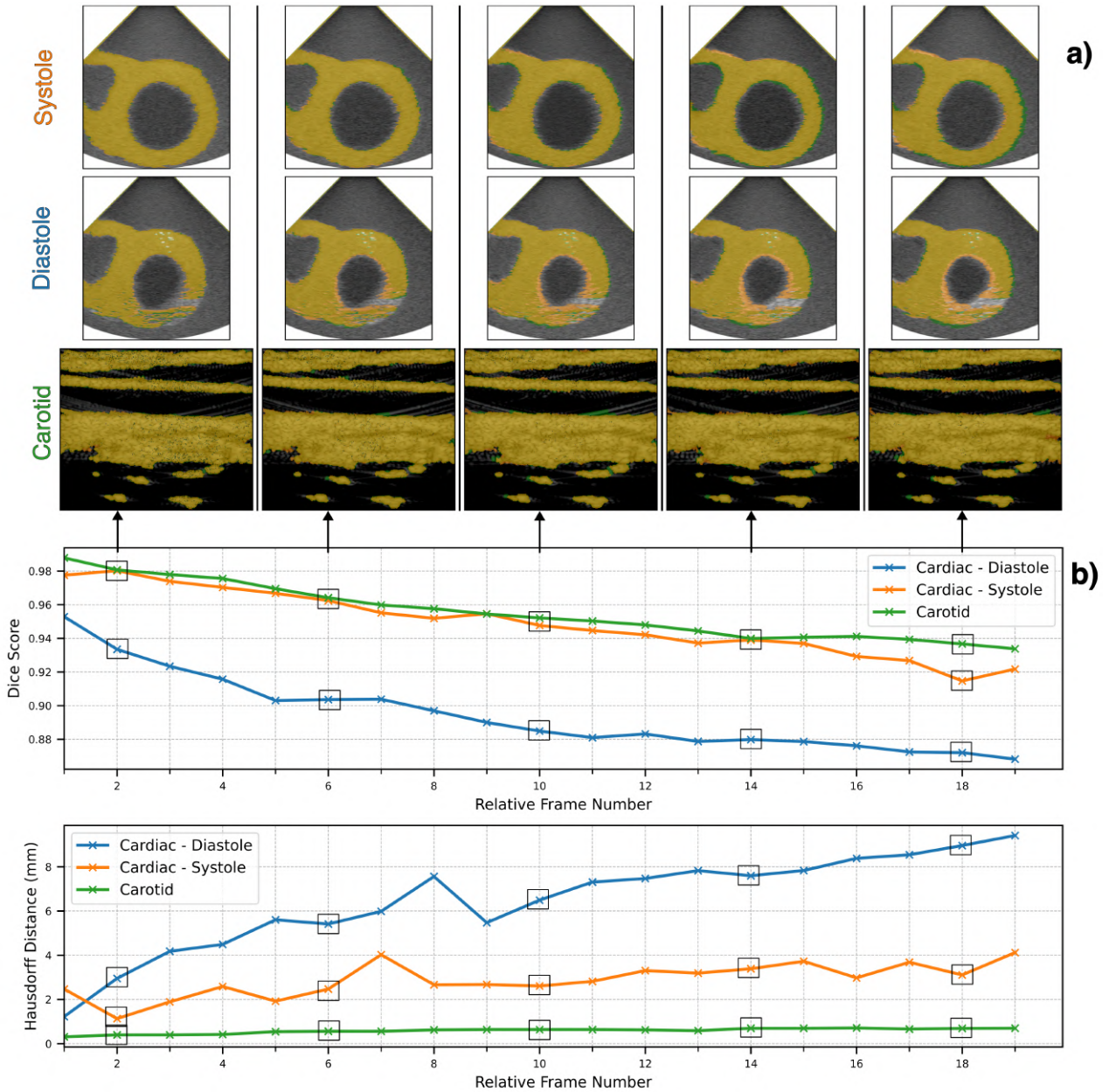
Figure 11: Results from SAS. (a) shows ground truth (green) and estimated (dark orange) masks overlayed to visualise mask alignment for frames 2, 6, 10, 14 and 18. The datatypes are cardiac systole (top), cardiac diastole (middle) and carotid (bottom). Yellow regions are indicative of areas of mask overlap. (b) The DS (top) and the HD (bottom) are plot against relative frame number and aligned with the respective segmentation masks from part a, shown by the black boxes. Legend shows cardiac diastole (blue), cardiac systole (orange) and carotid (green).

3.2.   Co-Registration

The effect of changing the co-registration parameters is given in Figure 12. α and β are positive and complementary, where $\alpha = 1 - \beta$, so an increase in β represents a larger weight for the masks in optimisation. The baseline method is conventional Dipy and is equivalent to a co-registration process where $\beta = 0$. The carotid dataset was selected due to its lower computation time, allowing for 25 systematically distributed parameter pairs to be evaluated.

For a description of the measures used, refer to section 2.4. Both the SSD and absolute error stay relatively constant until the critical value of 50% β. Here the SSD error begins to increase exponentially from the baseline whilst the absolute error decreases exponentially. The shaded regions represent the change in standard deviation of error from the baseline and show the same trends as the error metrics. For the carotid dataset, increasing to 100% β leads to an overall more spatially accurate registration with a lower mean (-17%) and standard deviation (-9%) of absolute error. Increasing to 100% β has the opposite effect on the intensity-based error. This increases the mean SSD by 153% and standard deviation by 16%. Overall, a decrease in absolute spatial error results in a large increase in intensity-based error.
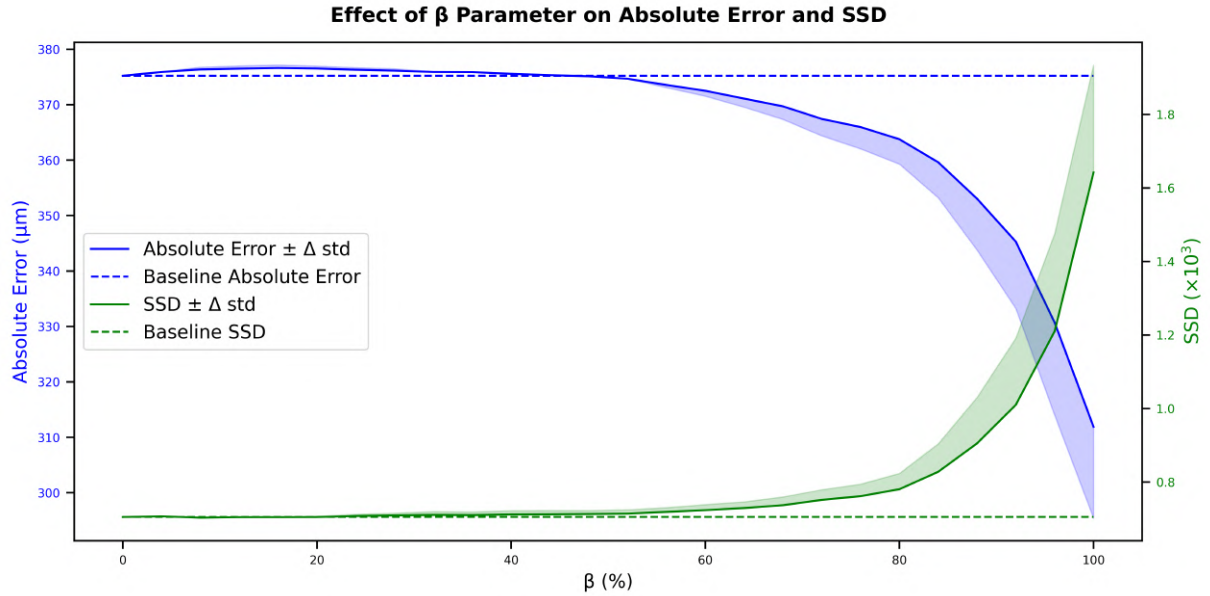

Figure 12: The effect of β parameter on absolute error (blue) and SSD (green). The data was collected from the carotid dataset using 25 systematically distributed parameter pairs. The baseline standard deviation of absolute error was 221μm. The baseline standard deviation of SSD was 390.

A comparative analysis of the performance between our baseline image registration method (Dipy) and our co-registration method for two datasets is shown in Figure 13. Parameter values of $\alpha = 0$, $\beta = 1$, were selected as this combination results in the smallest absolute error. A holistic evaluation of the algorithms was achieved using the following performance measures: intensity based (SSD, MI and NCC), anatomical alignment (absolute error and its standard deviation) and computational efficiency (reciprocal of execution time).

Co-registration demonstrates superior performance in terms of anatomical alignment with the absolute error and its standard deviation decreasing by 57% and 78% for cardiac (systole) dataset and by 17% and 23% for the carotid dataset. This represents a 132% and 20% increase in anatomical alignment accuracy for cardiac and carotid datasets respectively. The trade-off between spatial alignment and pixel alignment is exemplified by a 149% increase in SSD error, 28.5% decrease in MI and 3.3% decrease in NCC for the cardiac dataset. There were even larger increases in intensity-based error metrics for the carotid dataset. Importantly, the same trend for both datasets was observed; an increase in anatomical alignment, through co-registration, led to a reduction in traditional intensity-based similarity metrics. Additionally, the execution time for the co-registration algorithm was significantly reduced from the baseline: A 69.3% reduction (253.6 ± 48.5s to 77.9 ± 14.6s) for the cardiac dataset, and a 47.3% reduction for the carotid dataset (16.5 ± 5.4s to 8.7 ± 1.5s).
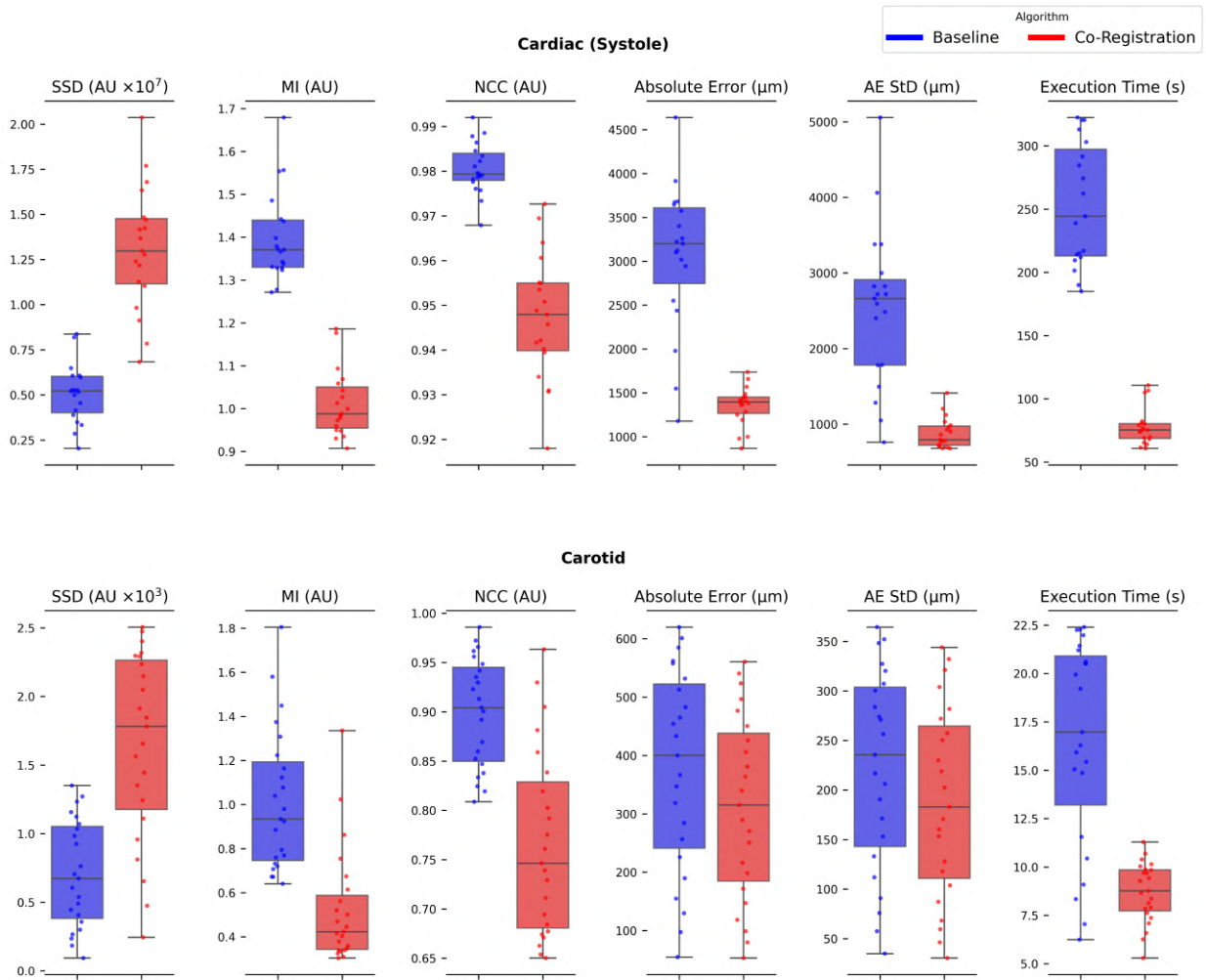
Figure 13: A boxplot compares the performance of co-registration (red) against the baseline method, Dipy (blue). The shaded area of the boxplot represents the interquartile range, with the median marked in the middle. The evaluation is shown for the cardiac systole dataset (top) and the carotid dataset (bottom) using the following metrics: SSD, MI, NCC, absolute error, standard deviation of absolute error, and execution time.

The Wilcoxon signed-rank test was used to test for statistically significant differences between the algorithms for two key metrics: execution time and absolute error. A Bonferroni correction was applied to account for multiple comparisons, adjusting the significance level to $\alpha\backslash 2$. The results of the Wilcoxon signed-rank test revealed a statistically significant difference in performance between the two algorithms at the 1% significance level, as seen in Table 3. The null hypothesis that there was no difference between the methods was rejected for all cases. Note here that the W-statistic is 0 as all datapoints saw a reduction in absolute error and execution time compared to the baseline.

Dataset	Metric	W	p	α (adjusted)	H₀ Rejected
Cardiac	Absolute Error	0.0	3.81e-06	0.005	√
	Execution Time	0.0	3.81e-06	0.005	√
Carotid	Absolute Error	0.0	2.38e-07	0.005	√
	Execution Time	0.0	2.38e-07	0.005	√

Table 3: The results from the Wilcoxon signed-rank test. Absolute error and execution time metrics were tested for significant difference for cardiac and carotid datasets. W represents the W-statistic, p the p-value and α the significance level.

Figure 14 (top) shows the estimated motion field from different methods. Figure 14 (middle) displays the absolute difference between the induced and estimated motion fields. Figure 14 (bottom) shows the output image for the different methods. Co-registration produces a smoother deformation field, creating a more realistic output image subject to smaller deformations.
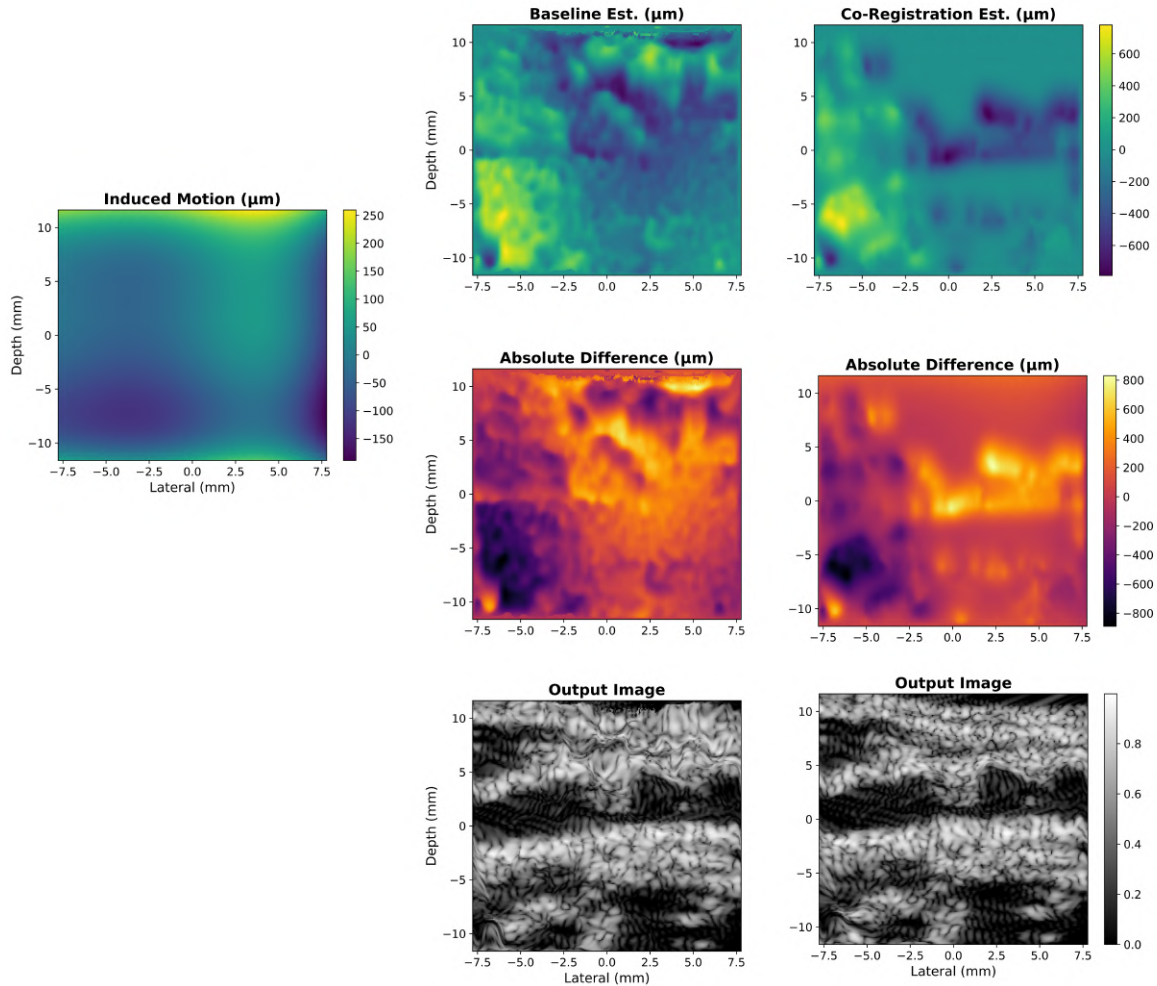

Figure 14: Left: simulated motion field applied to frame 20 of simulated B-Mode carotid images. Top: estimated deformation fields for baseline and co-registration algorithms. Middle: the absolute difference between the induced and estimated motion fields. Bottom: the final output image after deformation field has been applied. Inspired from [12].

4.   Discussion

4.1.   Semi-Automatic Segmentation

We used diffeomorphic image registration to segment ROIs in cardiac and carotid B-mode images. Overall, the results, seen in Figure 11, suggest that this method is suitable for simple segmentation tasks. The achieved alignment accuracy of < 95%, as measured by the DS, for both cardiac diastole and carotid datasets, is high enough to have confidence that our segmentations accurately model the underlying anatomy of the image. This value was much greater than that achieved by Sze and Tang in their SAS of the myocardium using diffeomorphic registration [11]. Although the level of deformation in the systole case was similar to that of Sze and Tang, we cannot directly compare these methods. The DS is often used to express segmentation accuracy in literature; however, the DS is highly shape-dependant and comparisons should be evaluated on a case-by-case basis. Large circular segmentation objects are likely to achieve a higher DS, particularly when their features initially overlap, such as in our case. Being small and distinctly shaped, the myocardium represents a difficult target for segmentation. Our large, rounded regions are bound to perform better in a spatial overlap measure, such as DS. To draw fair comparisons on accuracy, we would need to validate this segmentation method against a model dataset. As our method was similar to the method proposed by Sze and Tang, we would expect our method to be susceptible to the same limitations when tasked with complex segmentations, such as the myocardium.

The HD represents the displacement error of our estimated segmentation. The small HDs seen for cardiac (systole) maximum (4.12mm) and carotid (1.00mm) show that segmentation is a good model of the underlying anatomical structure of the image. Additionally, as expected, we see that larger deformations increase the maximum segmentation displacement. The relatively smooth increase in HD over the frames suggests a well regularised transform space with smooth realistic deformations; a poorly regularised deformation field would lead to regions of disproportionate deformation, resulting in larger HDs. Importantly, the HD quantifies true displacement error. In co-registration, using estimated masks from SAS, HD establishes the motion correction accuracy limit; the alignment accuracy cannot overcome the inherent alignment error of the initial masks. Hence for any future implementation of co-registration using our proposed pipeline, segmentation must be accurate to well beyond the required accuracy for motion correction. For SR US imaging of capillaries we would require an HD < 10μm [12]. In our study we were unable to achieve this level of segmentation accuracy.

Furthermore, the confidence in our results is limited by the accuracy of our ground truth measure. Our automatic segmentation method is prone to errors, as realised in the diastole segmentation from Figure 11a. These failed segmentations were due to a large deformation at the anterior left ventricle of the cardiac model, resulting in a distinct region of low intensity in the simulated B-Mode image. K-means clustering with only 2 clusters is unsuitable for segmenting multiple distinct regions [43]. As a result, we see a poor automatic segmentation for numerous frames during the diastole phase. SAS can often prove more accurate, as it only transforms existing mask regions. This makes it more robust to intensity changes compared to thresholding or clustering methods, which can introduce or remove regions between frames. Despite inaccuracies in our ground truth measure, the visual accuracy of both segmentation methods is sufficient to explore the benefits of using ROI masks in our co-registration method. However, the inaccuracy of ground truth measures limits our overall confidence in the estimated masks' calculated accuracy. This hinders our ability to draw definitive conclusions about the significance of masks in the co-registration process. Another limitation is that the same registration hyperparameters were used for all images pairs within a dataset. To improve the accuracy of our method, the number of pyramid levels, iterations and gaussian smoothing parameters could be optimised for each image pair, resulting in a more accurate registration and SAS.

4.2.   Co-Registration

Our novel co-registration method using the correct parameters lead to a significant decrease in the absolute error. By altering the co-registration optimisation parameters (α and β), we observed an inverse relationship between intensity-based error and absolute error metrics, given by Figure 12. This trade-off was expected. As α is decreased below 0.5 and β increased above 0.5, the optimisation is weighted towards the alignment of the masks rather than the images. The weight of image similarity in the optimisation is reduced, resulting in an increase in SSD, MI and NCC metrics. When optimisation is weighted towards mask alignment, the absolute error of the ground truth measure is reduced. This suggests that the mask is a better representation of the position of the underlying tissue than the intensities in the US image. This finding is remarkable in its significance; anatomical alignment accuracy of registration methods can be improved through registering accurate skeleton masks instead of the images. It suggests that mask-driven feature-based registration is more suitable for motion correction in US.

The co-registration approach with parameters: $\alpha = 0$, $\beta = 1$, was able to produce a 57% reduction in the absolute error, for realistic cardiac motion, while reducing the computational time by 69.3%. In comparison, Harput and Tang were able to reduce motion in simulated B-Mode images from 15.3μm, using only B-spline non-rigid registration, to 12.2μm using their two-stage method [12]. This represents a 20% reduction in the absolute error. They were also able to reduce computation time by 77.8%.

Both methods significantly reduce absolute error and computational time. Although our method achieved a greater reduction in absolute error, a direct comparison between the accuracy of the methods is invalid due to the vastly different levels of deformation used in the simulated images across both studies. However, comparisons can be drawn about the methodology of both studies. Harput and Tang first use affine registration to compensate for global motion, followed by non-rigid registration to compensate local deformation [12]. In contrast, our method first addresses local deformation by registering the images to produce a mask which is locally deformed to the image's distinct features. This is followed by a second registration, using the resultant mask, to estimate the global motion. This global motion step is more computationally efficient, which explains the decrease in the computational time for both algorithms.

In the case of co-registration, the input mask has a binary topology, so the global motion estimate is less susceptible to the surface variation of the underlying image, such as speckle. This brings about a key advantage of co-registration: its capability to create a smooth and realistic transform space. We see this in Figure 14, where the co-registration output image retains its natural morphology whereas the baseline method created harsh deformations and image blurring. Moreover, as given in Figure 13, the median and maximum standard deviation of absolute errors are substantially reduced with co-registration. This suggests that anatomical alignment has increased across all regions of the tissue. This contrasts with other registration approaches which naturally favour the alignment of regions of distinct intensity, such as conventional Dipy.

The largest limitation of the co-registration method is its reliance on accurate segmentation, as discussed in section 4.1, which forms the limit of accuracy for this method. Furthermore, the need to segment the image prior to co-registration increases the overall computation time. Moreover, this study is limited by the number of datatypes and degree of deformation evaluated. As a result, it is not possible to form assumptions on how this method would performs in complex registration tasks or with the small deformations expected within SR US imaging. Additionally, registration hyperparameters were optimised using the baseline method but used for all images and mask pairs within co-registration. These non-optimised parameters would likely hinder the algorithm's performance.

4.3.   Impacts on Motion Correction for SR US

The unified framework we developed uses a single image registration method for simultaneous image segmentation and motion correction. This has the potential to streamline the SR US imaging pipeline, providing ease of use and computational benefits. Unfortunately, due to computational challenges and the limited timeframe of this project, we were unable to test this framework on clinical data. However, our co-registration method shows promise in its clinical applications. Its substantial reduction in absolute error, up to 57%, and its capability to produce smooth and realistic deformations, making it a promising technique for motion correction in SR US imaging.

Importantly, this study has shown that traditional metrics for image registration such as SSD, MI and NCC are often inappropriate when only applied to the image. Image noise and interference effects can limit their accuracy when minimising anatomical displacement. We propose a semantic approach to image registration that evaluates both surface topology and positional anatomical estimates, such as masks, to construct a robust method for image registration.

4.4.   Future Work

Co-registration is currently limited to singular mask inputs. Future work could adapt this method to incorporate multiple masks allowing the optimisation to prioritise specific ROIs. This could increase the method's robustness to complex data. Additionally, two-stage motion correction could be incorporated into the co-registration methodology, potentially improving the overall accuracy of the method. Finally, the new insight into the significance of masks for accurate image registration could aid in the design of new techniques for motion correction.

5.   Conclusion

We use diffeomorphic image registration to develop a unified framework for image segmentation and motion correction for US images. This included establishing an effective method for ROI segmentation in a clinical setting and customising a best-in-class open-source image registration algorithm to simultaneously optimise mask and image alignment. This approach improved tissue spatial alignment by 132% in simulated cardiac data and reduced computation time by 69.3%. Although our results align with other motion correction methodologies, direct comparisons are limited due to differences in the evaluated images. Our work underscores the importance of using masks in image registration, highlighting their significant benefit for motion correction. This insight could pave the way for novel region-specific approaches to motion correction in SR US imaging.

6. References

[1] B. V. Zlokovic, 'Neurovascular pathways to neurodegeneration in Alzheimer's disease and other disorders', *Nat. Rev. Neurosci.*, vol. 12, no. 12, pp. 723–738, Nov. 2011, doi: 10.1038/nrn3114.

[2] J. Laitakari, V. Näyhä, and F. Stenbäck, 'Size, shape, structure, and direction of angiogenesis in laryngeal tumour development', *J. Clin. Pathol.*, vol. 57, no. 4, pp. 394–401, Apr. 2004, doi: 10.1136/jcp.2002.004978.

[3] R. K. Jain, 'Determinants of Tumor Blood Flow: A Review', *Cancer Res.*, vol. 48, no. 10, pp. 2641–2658, 1988.

[4] F. Lin, S. E. Shelton, D. Espíndola, J. D. Rojas, G. Pinton, and P. A. Dayton, '3-D Ultrasound Localization Microscopy for Identifying Microvascular Morphology Features of Tumor Angiogenesis at a Resolution Beyond the Diffraction Limit of Conventional Ultrasound', *Theranostics*, vol. 7, no. 1, pp. 196–204, Jan. 2017, doi: 10.7150/thno.16899.

[5] P. R. Moreno, K. R. Purushothaman, E. Zias, J. Sanz, and V. Fuster, 'Neovascularization in human atherosclerosis', *Curr. Mol. Med.*, vol. 6, no. 5, pp. 457–477, Aug. 2006, doi: 10.2174/156652406778018635.

[6] K. Christensen-Jeffries *et al.*, 'Super-resolution Ultrasound Imaging', *Ultrasound Med. Biol.*, vol. 46, no. 4, pp. 865–891, Apr. 2020, doi: 10.1016/j.ultrasmedbio.2019.11.013.

[7] B. Dollet, S. M. van der Meer, V. Garbin, N. de Jong, D. Lohse, and M. Versluis, 'Nonspherical Oscillations of Ultrasound Contrast Agent Microbubbles', *Ultrasound Med. Biol.*, vol. 34, no. 9, pp. 1465–1473, Sep. 2008, doi: 10.1016/j.ultrasmedbio.2008.01.020.

[8] O. M. Viessmann, R. J. Eckersley, K. Christensen-Jeffries, M. X. Tang, and C. Dunsby, 'Acoustic super-resolution with ultrasound and microbubbles', *Phys. Med. Biol.*, vol. 58, no. 18, p. 6447, Sep. 2013, doi: 10.1088/0031-9155/58/18/6447.

[9] C. Catana, 'MOTION CORRECTION OPTIONS IN PET/MRI', *Semin. Nucl. Med.*, vol. 45, no. 3, pp. 212–223, May 2015, doi: 10.1053/j.semnuclmed.2015.01.001.

[10] E. O.-T. Meng-Xing Tang, 'A Cost Effective and Open-Source Real-Time Respiration and ECG Phasing Device for Improved Ultrasound Imaging by Motion Correction'. Imperial College London - Department of Bioengineering, Apr. 2023.

[11] S. Sze, O. Bates, M. Toulemonde, M.-X. Tang, G. Bioh, and R. Senior, 'Semi-automatic Segmentation of the Myocardium in High-Frame Rate and Clinical Contrast Echocardiography Images', in *2022 IEEE International Ultrasonics Symposium (IUS)*, Oct. 2022, pp. 1–4. doi: 10.1109/IUS54386.2022.9958499.

[12] S. Harput *et al.*, 'Two-Stage Motion Correction for Super-Resolution Ultrasound Imaging in Human Lower Limb', *IEEE Trans. Ultrason. Ferroelectr. Freq. Control*, vol. 65, no. 5, pp. 803–814, May 2018, doi: 10.1109/TUFFC.2018.2824846.

[13] K. Knott *et al.*, 'The Prognostic Significance of Quantitative Myocardial Perfusion: An Artificial Intelligence Based Approach Using Perfusion Mapping', *Circulation*, vol. 141, Feb. 2020, doi: 10.1161/CIRCULATIONAHA.119.044666.

[14] J. R. L. F. Moccetti, 'Utilizing contrast echocardiography in practice', in *Essential Echocardiography: A Companion to Braunwald's Heart Disease.*, 2019, pp. 130–139.

[15] J. Ashburner, 'A fast diffeomorphic image registration algorithm', *NeuroImage*, vol. 38, no. 1, pp. 95–113, Oct. 2007, doi: 10.1016/j.neuroimage.2007.07.007.

[16] A. Trouvé, 'Diffeomorphisms Groups and Pattern Matching in Image Analysis', *Int. J. Comput. Vis.*, vol. 28, no. 3, pp. 213–221, Jul. 1998, doi: 10.1023/A:1008001603737.

[17] B. Avants, C. Epstein, and J. Gee, 'Geodesic Image Normalization in the Space of Diffeomorphisms', Oct. 2006.

[18] B. B. Avants, C. L. Epstein, M. Grossman, and J. C. Gee, 'Symmetric diffeomorphic image registration with cross-correlation: Evaluating automated labeling of elderly and neurodegenerative brain', *Med. Image Anal.*, vol. 12, no. 1, pp. 26–41, Feb. 2008, doi: 10.1016/j.media.2007.06.004.

[19] E. Garyfallidis *et al.*, 'Dipy, a library for the analysis of diffusion MRI data', *Front. Neuroinformatics*, vol. 8, p. 8, Feb. 2014, doi: 10.3389/fninf.2014.00008.

[20] Clara Rodrigo Gonzalez, 'An unpublished internal study'.

[21] A. Sotiras, C. Davatzikos, and N. Paragios, 'Deformable Medical Image Registration: A Survey', *IEEE Trans. Med. Imaging*, vol. 32, no. 7, pp. 1153–1190, Jul. 2013, doi: 10.1109/TMI.2013.2265603.

[22] I. Oezdemir, C. E. Wessner, C. M. Shaw, J. R. Eisenbrey, and K. Hoyt, 'Faster motion correction of clinical contrast-enhanced ultrasound imaging using deep learning', in *2020 IEEE International Ultrasonics Symposium (IUS)*, Sep. 2020, pp. 1–4. doi: 10.1109/IUS46767.2020.9251602.

[23] L. Bargsten and A. Schlaefer, 'SpeckleGAN: a generative adversarial network with an adaptive speckle layer to augment limited training data for ultrasound image processing', *Int. J. Comput. Assist. Radiol. Surg.*, vol. 15, no. 9, pp. 1427–1436, Sep. 2020, doi: 10.1007/s11548-020-02203-1.

[24] A. Ghaffari, R. Khorsandi, and E. Fatemizadeh, 'Landmark and intensity based image registration using free form deformation', in *2012 IEEE-EMBS Conference on Biomedical Engineering and Sciences*, Dec. 2012, pp. 768–771. doi: 10.1109/IECBES.2012.6498156.

[25] D. Mishra, S. Chaudhury, M. Sarkar, and A. S. Soin, 'Ultrasound Image Segmentation: A Deeply Supervised Network With Attention to Boundaries', *IEEE Trans. Biomed. Eng.*, vol. 66, no. 6, pp. 1637–1648, Jun. 2019, doi: 10.1109/TBME.2018.2877577.

[26] V. Bass, J. Mateos, I. M. Rosado-Mendez, and J. Márquez, 'Ultrasound image segmentation methods: A review', presented at the PROCEEDINGS OF THE XVI MEXICAN SYMPOSIUM ON MEDICAL PHYSICS, Merida, Mexico, 2021, p. 050018. doi: 10.1063/5.0051110.

[27] R. Prevost, R. Cuingnet, B. Mory, J.-M. Correas, L. D. Cohen, and R. Ardon, 'Joint Co-segmentation and Registration of 3D Ultrasound Images', in *Information Processing in Medical Imaging*, J. C. Gee, S. Joshi, K. M. Pohl, W. M. Wells, and L. Zöllei, Eds., in Lecture Notes in Computer Science. Berlin, Heidelberg: Springer, 2013, pp. 268–279. doi: 10.1007/978-3-642-38868-2_23.

[28] M. Ashfaq, N. Minallah, J. Frnda, and L. Behan, 'Multi-Modal Rigid Image Registration and Segmentation Using Multi-Stage Forward Path Regenerative Genetic Algorithm', *Symmetry*, vol. 14, no. 8, Art. no. 8, Aug. 2022, doi: 10.3390/sym14081506.

[29] J. M. Thijssen, 'Ultrasonic speckle formation, analysis and processing applied to tissue characterization', *Pattern Recognit. Lett.*, vol. 24, no. 4, pp. 659–675, Feb. 2003, doi: 10.1016/S0167-8655(02)00173-3.

[30] Y. Zhou *et al.*, 'A Framework for the Generation of Realistic Synthetic Cardiac Ultrasound and Magnetic Resonance Imaging Sequences From the Same Virtual Patients', *IEEE Trans. Med. Imaging*, vol. 37, no. 3, pp. 741–754, Mar. 2018, doi: 10.1109/TMI.2017.2708159.

[31] M. Lerendegui, K. Riemer, B. Wang, C. Dunsby, and M.-X. Tang, 'BUbble Flow Field: a Simulation Framework for Evaluating Ultrasound Localization Microscopy Algorithms'. arXiv, Nov. 01, 2022. Accessed: Nov. 21, 2023. [Online]. Available: http://arxiv.org/abs/2211.00754

[32] J. A. Jensen, 'Simulation of advanced ultrasound systems using Field II', in *2004 2nd IEEE International Symposium on Biomedical Imaging: Nano to Macro (IEEE Cat No. 04EX821)*, Apr. 2004, pp. 636-639 Vol. 1. doi: 10.1109/ISBI.2004.1398618.

[33] S. van der Walt *et al.*, 'scikit-image: image processing in Python', *PeerJ*, vol. 2, p. e453, 2014, doi: 10.7717/peerj.453.

[34] 'Open Source Computer Vison Library'. [Online]. Available: https://github.com/opencv

[35] B. B. Avants, N. Tustison, and H. Johnson, 'Advanced Normalization Tools (ANTS)'.

[36] P. Virtanen *et al.*, 'SciPy 1.0: fundamental algorithms for scientific computing in Python', *Nat. Methods*, vol. 17, no. 3, pp. 261–272, Mar. 2020, doi: 10.1038/s41592-019-0686-2.

[37] K. H. Zou *et al.*, 'Statistical Validation of Image Segmentation Quality Based on a Spatial Overlap Index', *Acad. Radiol.*, vol. 11, no. 2, pp. 178–189, Feb. 2004, doi: 10.1016/S1076-6332(03)00671-8.

[38] 'Hausdorff Distance - an overview | ScienceDirect Topics'. Accessed: Jun. 09, 2024. [Online]. Available: https://www.sciencedirect.com/topics/computer-science/hausdorff-distance

[39] T. Vercauteren, X. Pennec, A. Perchant, and N. Ayache, 'Diffeomorphic demons: efficient non-parametric image registration', *NeuroImage*, vol. 45, no. 1 Suppl, pp. S61-72, Mar. 2009, doi: 10.1016/j.neuroimage.2008.10.040.

[40] J. P. Thirion, 'Image matching as a diffusion process: an analogy with Maxwell's demons', *Med. Image Anal.*, vol. 2, no. 3, pp. 243–260, Sep. 1998, doi: 10.1016/s1361-8415(98)80022-4.

[41] 'Mutual Information - an overview | ScienceDirect Topics'. Accessed: Jun. 09, 2024. [Online]. Available: https://www.sciencedirect.com/topics/engineering/mutual-information

[42] 'normalized cross correlation - an overview | ScienceDirect Topics'. Accessed: Jun. 09, 2024. [Online]. Available: https://www.sciencedirect.com/topics/computer-science/normalized-cross-correlation

[43] N. Dhanachandra, K. Manglem, and Y. J. Chanu, 'Image Segmentation Using K -means Clustering Algorithm and Subtractive Clustering Algorithm', *Procedia Comput. Sci.*, vol. 54, pp. 764–771, Jan. 2015, doi: 10.1016/j.procs.2015.06.090.

7.   Appendix

7.1.   Code Repository

The algorithms and analysis scripts used in this study are available in the GitHub repository at https://github.com/elliot-tunstall/masked-imreg. The repository includes documentation and instructions for reproducing the results presented in this thesis.