scientific reports

OPEN

# A concept for fully automated segmentation of bone in ultrasound imaging

Ananth Hari Ramakrishnan[1], Muthaiah Rajappa[1], Kannan Krithivasan[1], Nachiappan Chockalingam[2], Panagiotis E. Chatzistergos[2,4] & Rengarajan Amirtharajan[3]✉

**This study proposes a novel concept for the automated and computerised segmentation of ultrasound images of bone based on motion information. Force is applied on the heel region using the ultrasound probe and then removed while recording the video of the bone using ultrasound. The interface between the bone and surrounding tissues is the region that moves with maximum speed. This concept is utilised to determine a map of movement, where speed is the criterion used for the bone segmentation from the surrounding tissues. To achieve that, the image is subdivided into regions of uniform sizes, followed by tracking individual regions in the successive frames of the video using an optical flow algorithm. The average movement speed is calculated for the regions. Then, the regions with the higher speed are identified as bone surfaces. It is given as the initial contour for the Chan–Vese algorithm to achieve smoother bone surfaces. Then, the final output from the Chan–Vese is post-processed using a boundary tracing algorithm to get the last automated bone segmented output. The segmented outcomes are compared against the manually segmented images from the experts to determine the accuracy. Bhattacharyya distances are used to calculate the accuracy of the algorithmic and manual output. The quantitative results from Bhattacharyya distances indicated an excellent overlap between algorithmic and manual works (average ± STDEV Bhattacharyya distance: 0.06285 ± 0.002051). The bone-segmented output from the optical flow algorithm is compared with the model output and the texture-based segmentation method's output. The work from the motion estimation methods has better segmentation accuracy than the model and texture segmentation methods. The results of this study suggest that this method is the first attempt to segment the heel bone from the ultrasound image using motion information.**

**Keywords** Heel bone, Map of movement, Segmentation algorithm, Optical-flow, Chan–Vese

Bones, as the essential components of the skeletal system, work alongside tendons, ligaments, and muscles to form the musculoskeletal tissues. Bones are protective shields for internal organs and are crucial in safeguarding red and white blood cells. Common imaging modalities for bone include CT, MRI, and ultrasound. CT is typically used in pre-operative care and intervention planning[1] but comes with the drawback of ionising radiation, limiting its frequent use. MRI offers a high-resolution alternative for fracture investigation but is expensive and lacks portability. Ultrasound, favoured for its low cost, real-time imaging, and non-invasive nature, is increasingly preferred over other imaging modalities due to its non-ionising radiation and portability. However, low contrast, noise, artifacts, and high inter- and intra-observer variability make manual ultrasound image analysis tedious and unreliable, underscoring the need for robust automated segmentation methods[2].

Ultrasound is particularly beneficial for bone imaging because of its radiation-free[3], real-time capabilities, and use in guiding non-surgical procedures like epidural anaesthesia and spinal blocks[4]. Despite its advantages, speckle noise and imaging artifacts have limited its application as a standalone tool for intraoperative imaging. To enhance image guidance in orthopaedic procedures, registering ultrasound images with MRI/CT is critical, with bone segmentation from ultrasound images being a key preliminary step. When ultrasound signals encounter the region between two tissue interfaces, the resulting image features high-intensity pixels due to impedance mismatches. Bone, having the highest impedance among tissues, appears as a high-intensity region, followed by a low-intensity "shadow" region characterised by artifacts, as the interior of the bone is not imaged. This study

[1]School of Computing, SASTRA Deemed University, Thanjavur 613401, India. [2]Centre for Biomechanics and Rehabilitation Technologies, School of Health, Science and Wellbeing, Staffordshire University, Stoke-on-Trent ST4 2DE, UK. [3]School of Electrical & Electronics Engineering, SASTRA Deemed University, Thanjavur 613401, India. [4] Biomedical Engineering, University of Dundee, Dundee, Scotland, United Kingdom. ✉email: amir@ece.sastra.edu

focuses on the calcaneus (heel bone) of the hindfoot. In ultrasound analysis of this region, the high-intensity bone boundary is observed adjacent to the macrochamber layer, a globular fat layer in close contact with the bone.

In contrast, the superficial microchamber layer lies just above it. Segmenting the heel bone from ultrasound images enables the analysis of the calcaneus's mechanical properties, offering valuable insights for predicting diabetic foot ulcers[5]. Additionally, calcaneus segmentation is utilised to diagnose and predict conditions such as plantar fasciitis, arthritis[6,7], and osteoporosis[8]. The non-ionising nature of ultrasound makes segmentation particularly beneficial for evaluating bone health in children[4].

Automated bone segmentation serves as a complementary tool for radiologists, assisting in diagnostic tasks and facilitating visualisation. Accurate bone segmentation aids in pathology localisation and treatment planning. This research aims to enhance the computerised segmentation of hard tissues by promoting further exploration of movement mapping techniques. For the first time, this study demonstrates that hard tissues can be accurately segmented using movement-based information. Ultrasonography has shown high diagnostic accuracy in detecting upper and lower limb bone fractures in adults, particularly in foot and ankle fractures, where bone segmentation is a crucial initial step. Further studies involving a larger number of participants and other anatomical regions are necessary to validate and strengthen the quality of existing evidence before recommending ultrasound as a primary imaging modality for peri-operative applications.

The most common methodology used in bone segmentation is Snake models[9–12], Geodesic active contours[13], Log-Gabor wavelets[14–18], Random Forest[19], locally statistical level set methods[20], fully convolutional networks[21], Convolutional neural networks[22–24], BoneNet CNN[25,26], dual-task ultrasound transverse vertebrae seg-mentation network[27]. Even though these methods are unique, they all depend on the same type of input information and try to distinguish the region of interest based on an analysis of intensity, texture, and shape features. However, in the case of ultrasound imaging of bone, such information might not be sufficient to delineate the hard tissue accurately. In such cases, the movement pattern could help with boundary perception and accurately segment bone.

Motion information was incorporated earlier for the computerised segmentation of ultrasound images of connective tissues (soft tissues)[28]. Observing how the image changes with probe movement is key for detecting tissues of interest[9,10]. In[29], the force is applied to the soft tissue to enhance the bone boundary in the power doppler images. The difference in acoustic properties between bone and soft tissues in power Doppler imaging aids in distinguishing the bone from surrounding tissues, even in the presence of noise. Hence, this work[29] helps enhance bone surfaces and facilitate bone segmentation. From the literature on bone segmentation from ultrasound images[30], no automated segmentation method performs segmentation using motion information. Hence, this concept can be used for the automated segmentation of bone from B-mode Ultrasound Images.

The proposed method introduces a novel algorithmic framework for automated segmentation based on motion information. To create a movement pattern, force is applied to the heel pad using the ultrasound probe, and then force is removed while recording the ultrasound video.

The highlights of this work are

- Motion (optical flow) based heel bone segmentation from the ultrasound image is done
- Different regions in the image are tracked in the successive frame using an optical flow algorithm, and the centroid of the regions is saved. Strain and average movement speed is calculated from the centroid values.
- Out of these two methods, average movement speed better differentiates the bone from the surrounding tissues.
- From the highest average movement speed values from each box segment, the bone region coordinates, which move at a higher speed than surrounding tissues, are extracted and given as initial contour to the Chan–Vese segmentation.
- The final segmented output is compared with the manually segmented image to compute the segmentation accuracy. Bhattacharyya distance is used for accuracy computation.
- The Bhattacharyya distance between the manual and algorithmic output is 0.0643 for participant 1 (almost close to 0) and 0.0614 for participant 2, showing good quantitative agreement between the two outcomes. The lower the Bhattacharyya distances, the higher the accuracy between the outputs.
- These results were compared with the existing model and texture-based methods. The Bhattacharyya distances for model and texture-based methods are 1.0927 and 5.3961 for Participant 1 and 1.5233 and 0.55807 for Participant 2. The comparison results indicate that Bhattacharyya distances are lower for our proposed approach.
- The time complexity of the overall algorithm is O[(K + F) N], where K is the number of frames, F is the features, and N is the quantity of pixels on the frame.

In this context, this paper proposes a concept of bone segmentation based on motion information. A brief presentation of existing methods previously used for bone has been presented, and then existing motion tracking algorithms are presented in Section "Related works". The optical flow will be employed to track different regions in bone, and then the average movement speed will be calculated based on the displacement of the regions. Since bone moves at a higher speed than the surrounding regions, the bone region is differentiated from the rest of the regions and finally segmented. Section "Concept demonstration" describes the optical flow algorithm for tracking and strain VS average movement speed. Section "Proposed methodolog" describes detailed procedures and parameters regarding the optical flow and the appropriate figures. In Section "Results", results from the bone segmentation are described qualitatively and quantitively by comparison against the manual segmentation. The work's advantages, drawbacks, and future directions are discussed in Section "Discussion".

# Related works

## Existing segmentation methodologies in Bone segmentation

A structured review of the literature on the segmentation of bone using PubMed revealed twenty-four papers. Existing algorithms for bone segmentation from ultrasound images are as follows. Segmentation of long bones in infants[4] uses the thresholding method. This involves two steps: outer skin contour extraction, followed by inner bone segmentation. In skin contour extraction, thresholding is done in an iterative process until a good contour is obtained. Bone surface extraction is also done by thresholding, where threshold selection is the solution to the depreciation of $|gz(t) - Sz|$, where $gz(t)$ is the histogram of the zone, and Sz is the surface estimation. The iterative skin contour extraction makes the whole process semi-automated.

Segmentation of Calcaneus bone[9], ulna, clavicle, humerus[10], metacarpal and phalanges bone[18], and synovial boundaries[24] are done using the snake contour method[31]. In calcaneus bone segmentation, parametrisation of an initial contour based on the Fourier descriptor to rectify the initial contour initialisation drawback in the snake model should be as close as possible to the region of interest. Fourier descriptors are calculated from manually segmented calcaneus images, making the segmentation process semi-automated[9]. The femoral head and acetabulum are segmented using Geodesic active contours[23], where energy minimisation is based on gray level and geometric features. The geometric features are obtained from Shape-priors, making it a semi-automated method[32].

The lamina detection from an ultrasound image using Log-Gabor filters extracts the image's phase features (ridge structures). The ultrasound image is cross-correlated with the lamina template using Pearson cross-correlation. Then, the points of highest cross-correlation are marked, and the area around the point is zero. There are three laminas present between the skin and ligamentum flavum. Hence, the 2nd and 3rd highest correlation points are computed[14]. The above segmentation process is also done semi-automated. Random forest classifier, which comes under the machine learning method, classifies and extracts the spinous process bone surface from the shadow feature, thus classifying it as bone and non-bone pixel[19]. Segmentation of cartilage thickness is computed using Locally Statistical Level Set methods[33,34] (LSLSM).

The energy function in LSLSM comprises global similarity, local statistical, and penalty terms. The local statistical term overcomes the inhomogeneity problem by constructing circular regions around the pixel and calculating average pixel filtering, which is transferred into a new image domain. The global similarity term minimises the Bhattacharyya coefficient between the probability distribution function of intensities inside and outside the contour. The penalty term is added to maintain the signed distance property of the contour at the zero-level set. This method may fail if the background intensity distribution is similar to the foreground. Radius and tibia bones are segmented from ultrasound images using BoneNet CNN (Modified U-net) architecture[25]. Bones in synovial joints is segmented from ultrasound images using a Convolutional neural network[24]. Transverse vertebrae segmentation from ultrasound images is achieved using a dual-task ultrasound transverse vertebrae segmentation network[27]. Here, two separate U-net streams are used for semantic and boundary segmentation.

As mentioned earlier, the algorithms used for bone image segmentation are mainly semi-automated. Our ultimate aim of this work is to develop a fully automated method for the segmentation of bone. Also, there are other drawbacks to the existing methods. The active contour/snake contour method uses gradient information for contour deformation, which is unreliable in ultrasound images due to speckle noise and image artifacts[35]. In the Random Forest[1]classifier, the presence of many trees makes the algorithm slow. Generally, these algorithms are quick to train but take a long time to predict. Also, machine learning and deep learning-based approaches require training in manually segmented images in large numbers[36,37], which is difficult due to speckle noise and low contrast in ultrasound images[38]. Also, it is challenging to create a dataset comprising all physiological and anatomical variabilities among different subjects, which is a complicated and tedious task[39].

This research article is the first to implement a tracking algorithm for automated bone segmentation, utilising motion information from different regions of the image tracked across video frames to distinguish the bone region from the surrounding tissue. This study combines existing semi-automated segmentation methods to develop a fully automated segmentation approach that complements model-based and texture-based segmentation techniques.

This paper proposes incorporating motion information to automate the segmentation process for accurate bone identification. The method involves tracking uniform-sized regions in the images using an optical flow algorithm and calculating strain and average movement speed. The bone is distinguished from surrounding regions based on its higher movement speed relative to the surrounding areas in the video.

Different regions in the video frames are tracked using a motion estimation algorithm. Selecting a motion estimation algorithm that can reliably track the other regions in ultrasound images is essential. The movement of structures can be estimated by finding the spatial correspondences between frames. Initially, motion estimation is accomplished using a traditional image registration algorithm, where the spatial transformation is done by aligning the successive frame (i.e., the frame that has moved) with the initial stationary frame. The spatial transformation used in this method is affine transformation, which maps the points from one frame to another. This transformation can map scaling, shear and rotations in the image plane. Tracking algorithms must be fast and consistent for fast-moving videos and videos of large time intervals. Lucas and Kanade[40] calculated the match between two blocks in successive frames using an image registration technique that deploys the spatial intensity gradient of the image. It employs fewer and potential matches between the images and can handle rotation, scaling and shearing. When generalising the one-dimensional technique to higher dimensions, linear approximations are used, which introduces errors, which is the drawback of this method.

## Motion estimation methods

The motion estimation methods can be broadly classified into intensity and feature-based methods[40]. In intensity-based methods, the spatial transformation of the pixels in the successive frame is based on the pixel intensity, which is optimised by iterative methods to improve the similarity criterion. In feature-based methods, features based on edge, texture or shape are extracted from the image, and spatial correspondences are found in the successive frame. Our application aims to segment the bone region based on displacement; hence, feature-based methods are unsuitable for our application. Various techniques to compute the displacement between two frames are the sum of absolute differences[41], the sum of squared differences[42], correlation coefficient[43], normalised cross-correlation[44], and optical flow algorithm[45].

There are deep learning-based methods for motion estimation. They can be classified into supervised convolutional neural networks[46–48] and Unsupervised convolutional neural networks[49].

Bohs and Trahey[41] tracked the speckle patterns in moving blood. Non-overlapping kernel regions are tracked in successive images using the sum of absolute differences algorithm (SAD), which uses non-normalised absolute error for pattern matching. The method requires one absolute difference operation.

Paramkusam and Reddy[42] tracked video sequences using the sum of squared differences. Here, the similarity measure is based on the intensity differences between the images, which is assumed to follow a Gaussian distribution (additive noise)[50]. However, ultrasound images are characterised by speckle noise (multiplicative noise). Hence, the sum of squared differences is unsuitable for tracking the ultrasound images.

Golemati et al.[43] tracked the carotid artery wall motion in ultrasound video sequences using correlation coefficient as a matching criterion. This block-matching technique can compensate for the local variations in the mean and standard deviation of the signals. The region of interest (ROI) size affects the motion tracking. Hence, optimal ROI size is selected to accurately monitor the region of interest. The higher computational cost is the drawback of this method.

Chernak and Thelen[44] tracked the tendon region in an ultrasound video using normalised cross-correlation. The kernel window is initialised in the first frame and a search window of size greater than the kernel window is initialised around the same neighbourhood in the second frame. When normalised cross-correlation is calculated between kernel and search windows, the point at which maximum correlation occurs is the highest probability of occurrence of kernel window. This is where the kernel window is displaced in the subsequent frame. Thus, normalised cross-correlation can help compute the motion of tendon tissues in subsequent frames under loading conditions. Normalised cross-correlation is most commonly used in ultrasound elastography applications[51–53]. Speckle noise variations affect the tracking results, which is the drawback of this approach[54]. Korstanje et al.[51] proposed a multi-kernel block-matching system to overcome this drawback. The reference block is divided into several sub-blocks, then block-matching is computed for each subblock to find the closest match. The matching results are consolidated to arrive at the overall results. Normalised cross-correlation is the block-matching method used here. However, the drawback of this approach is poor performance if the tracking target's motion is minimal[54]. The optical flow algorithm estimates the motion vectors of pixels between frames. In contrast, the rest of the techniques (sum of absolute differences, the sum of squared differences, correlation coefficient and normalised cross-correlation) generate a similarity score representing how close the regions/pixels are to each other. Optical flow can effectively handle scaling, rotation, and non-rigid transformations, whereas the different techniques cannot accommodate these changes.

The optical flow algorithm is used to track the motion of the needle head[55], left ventricle[56]and muscle aspect ratio in the rectus femoris (RF) muscle[57]. Optical flow estimates the spatiotemporal change in pixel intensities between the frames to calculate the velocity of the pixels[18]. The optical flow algorithm assumes that the pixel intensity is constant between the frames. Displacement is not adequately estimated if the motion is too large, which is the drawback of this method[58]. This drawback is rectified by using image pyramids[59].

The architectures employed for supervised convolutional neural networks include the Siamese tracker (SiamFC), FlowNet2, and Cascaded One-shot Deformable Convolutional Neural Network (OSD-CNN). As supervised methods, they require training on labelled datasets, which depend on large collections of manually segmented images that are currently unavailable.

MaskFlownet-based unsupervised convolutional neural network (MF-UCNN) is commonly used for unsupervised convolutional neural networks. MF-UCNN relies on loss functions to guide network training, utilising image similarity and the smoothness of axial and lateral displacement as loss metrics in ultrasound imaging. However, ultrasound images are inherently noisy, and loss functions based on image similarity often fail to account for these noises, resulting in inaccurate motion estimation[60,61].

Drazan et al.[45] tracked the muscle fascicle in the successive frames of the ultrasound video using an optical flow algorithm. In this approach, the Kanade-Lucas-Tomasi (KLT) algorithm[40,45,62] is used to determine the object movement between the frames. This algorithm uses image pyramids[59] to handle the large displacement of pixels between the frames efficiently. Affine transform is obtained using optical flow, which tracks the regions in successive frames. Optical flow performs displacement computation at multiple levels of resolution in image pyramids, where each level is reduced by a factor of two. The tracker starts from the lowest resolution level and continues tracking to the higher resolution levels till convergence.

The commonly used methodologies for bone segmentation in ultrasound images are model-based, machine-learning and deep learning-based methods[63–68]. Though these methods are unique, they all depend on texture[69], shape[70] and intensity information[71] to differentiate the zone of interest from the background. However, in the case of ultrasound imaging of bone, such information is not enough to accurately segment hard tissue's boundaries[72,73]. In such scenarios, video analytics, where motion patterns are considered, could improve the analysis and result in accurate segmentation.

## Concept demonstration

The Bone segmentation algorithm was implemented in the heel region of the foot. Two videos were collected from two human participants. These two videos were collected as a part of a bigger project. Relevant guidelines and regulations are carried out in all methods. Ethical approval was secured from the Staffordshire University

Ethics committee, Stoke-on-Trent, England. All participants provided written informed consent before any data was collected.

The anatomical structure of the calcaneus bone, as visualised in ultrasound imaging at the apex of the heel, is organised as follows. The region above the calcaneus bone comprises the fat pad and the skin. The fat pad has two distinct layers: a superficial microchamber layer (stiffer in nature) closer to the skin and a deeper macro chamber layer (more deformable) above the calcaneus bone[74]. The primary function of the fat pad is to absorb and distribute shock during locomotion by deforming under pressure. The mechanical behaviour of the fat pad differs under compression, with the microchamber layer being ten times stiffer than the macrochamber layer. As a result, when the fat pad is compressed, the macro chamber layer deforms more significantly, providing the cushioning effect of the heel pad[75]. This deformation creates relative movement between the bone and the skin, which is captured while recording the ultrasound video of the bone.

The external force to the plantar soft tissue is applied using an ultrasound probe[76]. The dynamic indentation test was performed using a motorised ultrasound indentation device consisting of a linear array ultrasound probe series with a load cell. The device performs linear movement in one direction using an actuator and controller. The indentation device records the applied force and displacement, while the clinical ultrasound unit records B-mode images in the frontal plane. The external force is applied to the plantar soft tissue to quantify the heel pad's mechanical properties, which helps analyse diabetic foot ulceration[5,77]. In the testing set-up used to collect the images, the heel bone is practically fixed in space and remains undeformable. Because the probe is the only component moving during testing, bone appears to be the fastest-moving object relative to the probe[77]. This concept can segment the fastest-moving bone from the slower-moving surrounding tissues in the ultrasound video. Hence, at the start of the video, the distance between the calcaneus and bone is more incredible, while in the compressed stage, the distance between the calcaneus and skin is less. Heel bone segmentation from ultrasound video is done in an automated way using the following steps.

Thirty load/unload cycles were performed in the dynamic indentation test, targeting a subject-specific force with a displacement rate of 21 mm/s. This process was conducted during three independent trials, each involving the participant standing still for 10 s[77]. Two videos were recorded by two participants, with image dimensions of 652 pixels in height and 800 pixels in width. The frame rate for the first participant's video was 28 frames per second, while the frame rate for the second participant's video was 25 frames per second.

The block diagram for the overall work is given in Fig. 1. The classical method for bone delineation relies on manual segmentation, which is time-consuming, labour-intensive, and subjective. To address these limitations, computerised segmentation from static ultrasound images has been introduced to accelerate the process. However, detecting bone boundaries in static images is challenging due to the similarity between bone and surrounding tissues. To overcome this, motion information is utilised for bone segmentation. The image is divided into distinct regions, and each region is tracked using an optical flow-based motion estimation algorithm to calculate the centroid in each frame. The average movement speed across each column is then computed, and the column with the maximum speed is identified as corresponding to the bone region. The bone centroid pixels are connected to delineate the bone region, which is subsequently refined using the Chan–Vese segmentation algorithm.

## Proposed methodology

### Optical flow-based region tracking

To automate the bone segmentation process in ultrasound, this work proposes to include motion information in the segmentation of bone. Instead of analysing a single image, motion-based segmentation analyses a series of video images to segment the region of interest from the background. The following steps have been carried out to accomplish the task.

*Step 1*: A Box segment is drawn vertically in the right corner of the image, covering the bone and skin region.

*Step 2*: The rectangular boxes (to be tracked) are automatically filled all over the image based on the height, width, horizontal and vertical overlap.

*Step 3*: The tracking process is accomplished using the pointTracker function in MATLAB. The movement of interest points between two frames in a video is determined using an optical flow algorithm.

*Step 4*: Tracking of KAZE (Box tracking between consecutive video frames is performed using the sparse optical flow algorithm called Lucas-Kanade method.) interest points and calculation of 2D geometric transformation matrix between two frames are selected by trial-and-error method and kept constant throughout the analysis. In Sparse optical flow, 'features' characterising the box are tracked rather than the individual pixels (dense optical flow method). The box's features can include corners, edges, blobs, multiscale blobs, or regions of uniform intensity. This work uses the multi-scale blob, known as KAZE interest features, to characterise the box. The Lucas-Kanade method tracks These KAZE features between consecutive frames in the video.

*Step 4.1*: This function incorporates the KLT algorithm to compute a region's optical flow and assess the movement of the region of interest between a pair of frames.

*Step 4.2*: The hyperparameters used in this function are (1) Block size (2) Number of levels in image pyramids (3) Maximal bidirectional error & (4) iterations needed to obtain the optimum solution. Block size, number of levels in the image pyramid, maximal bidirectional error and maximum number of iterations were set to 5 mm × 5 mm, 9, 3 mm and 40 throughout the experiment.

*Step 4.3*: This function output gives the new location of KAZE interest points in the successive frame. The number of octaves in optical flow-based tracking should be an integer greater than 1 to enable multiscale analysis.

*Step 4.4*: More features could be detected when the number of octaves is more significant. Similarly, the number of scale levels is an integer in the range[14,16], and smooth scale changes could be achieved if number of scale levels is high.

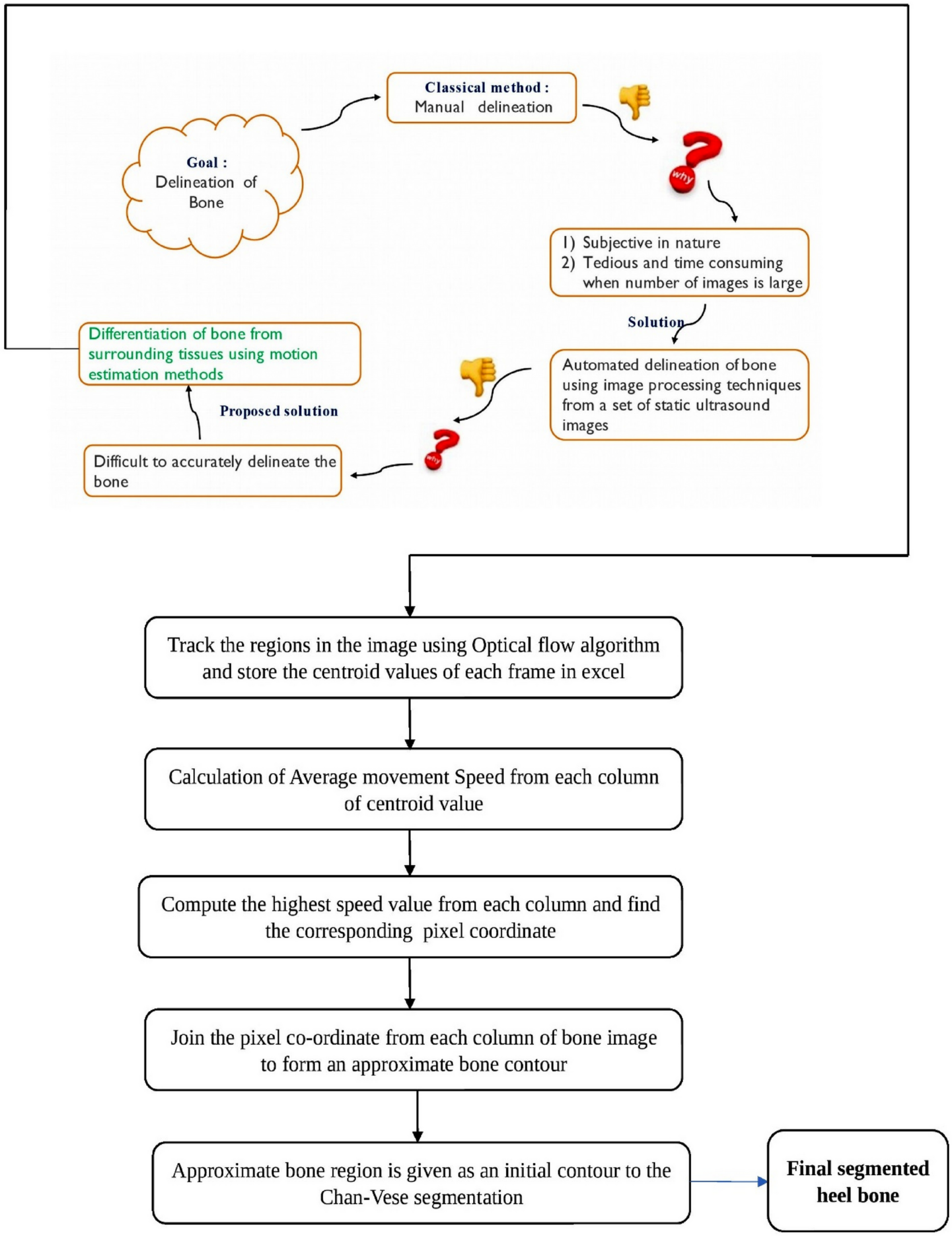


**Fig. 1**. Block diagram of overall work.

*Step 5*: MATLAB determines translation, rotation, scaling, and distortion of the region of interest using the estimateGeometricTransfrom2D function.

*Step 5.1*: This is done by inlier mapping present in the KAZE interest points of the first frame to the KAZE interest points of the successive tracked frame.

*Step 6*: The box's position in the next frame is found by applying the affine matrix given by the estimateGeometricTransform2D function.

*Step 7*: After tracking the first frame, the tracked frame becomes the reference frame for the successive frame, and then the process is repeated for all the frames in the video.

The boxes are tracked, and the centroid of the boxes successfully tracked in all frames is stored in Excel sheets. Each column represents the box coordinates in successive frames of the video. The successive Box segments are stored in separate Excel sheets. The displacement of the boxes between successive frames, referred to as individual box speed, is determined by calculating the difference between the current and previous columns. The average displacement is obtained by taking the mean of the individual displacements across successive frames in the video. The average movement speed is then computed as the ratio of the average displacement to time. Since the time interval between successive video frames is constant and equal to 1, the average movement speed equals the average displacement of the boxes. Since bone moves with the highest speed compared to the surrounding tissues, coordinates corresponding to the highest value from the average movement speed give the bone region approximately. To obtain the smooth surface of the bone, the coordinates are given as the initial contour for Chan–Vese segmentation[78].

## Strain versus average movement speed calculation for bone segmentation

The Box segment is constructed to cover both the bone and skin regions, and boxes are generated automatically for one Box segment alone based on user inputs such as height, width, and overlap of the boxes horizontally and vertically, as shown in Fig. 2. The boxes are tracked automatically using an optical flow tracking algorithm, and the x and y coordinates of the centroid values are stored in sheet 1 of Excel sheet (y-coordinates -Box1-26 files, x-coordinates X1-26 files, present in Supplementary material B). From the centroid values, we aim to differentiate the coordinates of the bone from the surrounding regions. The bone moves vertically in the ultrasound video. Hence, bone moves on the y-axis, and there is no change in bone movement along the x-axis. Therefore, the y-coordinate of the centroids is processed for further analysis. Two parameters are calculated to differentiate the bone from surrounding regions: Strain and average movement speed. The force is applied at the heel region through the ultrasound probe; skin regions get deformed throughout the video of the bone, whereas bone regions are un-deformed throughout the video of the bone. Hence, strain is calculated for springs (Box segment between the centroid of two boxes in a Box segment; if n boxes are present, n-1 springs are generated) between the last frame and first frame of the video using the formula (final deformed length – original deformed length / original deformed length).

The plotted strain between springs and strain is shown in Fig. 3 for Participants 1 and 2. The average movement speed is the average displacement of the boxes tracked between the frames divided by the time. In the ultrasound video of the bone, the bone moves at the highest speed compared to the surrounding regions. Hence, the box with the highest average movement speed corresponds to the bone region, as shown in Fig. 4. In Fig. 3, the difference between the bone and the surrounding region is unclear, whereas in Fig. 4, there is a clear-cut difference between the bone and skin regions. Hence, the average movement speed is deployed to differentiate the bone from the skin regions. The boxes are generated for all the box segments in the image, and then the bone centroid from each box segment is extracted using the highest average movement speed. Then, the bone centroids are joined together to get the approximate bone region, which acts as an initial contour to the Chan–Vese segmentation for smoother bone extraction in the next step.

## Average movement speed-based bone characterisation from ultrasound images

From Section "Strain versus average movement speed calculation for bone segmentation", it is concluded that average movement speed is better for characterisation from bone. In this section, we will deploy the average movement speed for bone segmentation.

In Step 1, a Box segment is drawn at the right end of the image. Then, the user gives the width and height values of the box along with horizontal and vertical overlap. Based on these inputs, the boxes are automatically filled in the image. The width and height of the boxes are given as 30 and 10 pixels based on the trial-and-error method (the lowest possible size of the box that can be accurately tracked). The Box segment initialisation and box initialisation are shown in Fig. 5.

In step 2, the boxes are tracked in successive video frames using the KLT algorithm. The centroid of the boxes tracked in all the video frames is stored in Excel sheets, and the remaining boxes lost before reaching the last

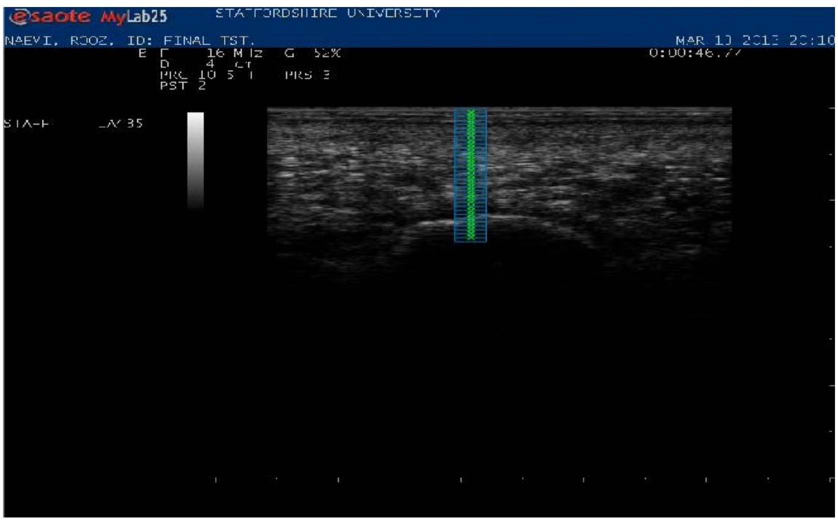

**Fig. 2**. Box segment covering bone and skin region.

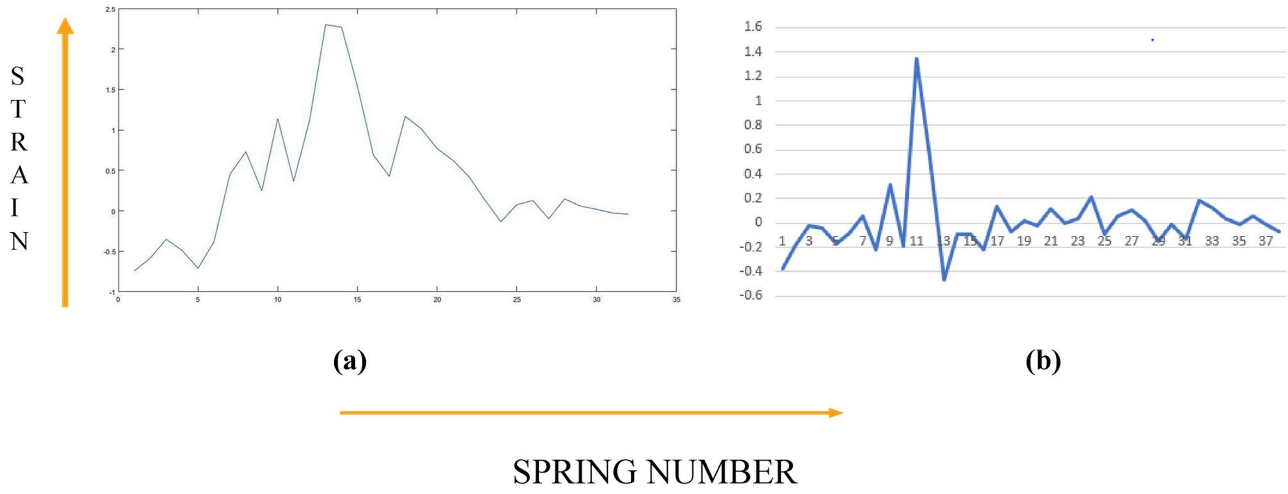


**Fig. 3**. Strain map between spring number and strain for (**a**) Participant 1 and (**b**) Participant 2.

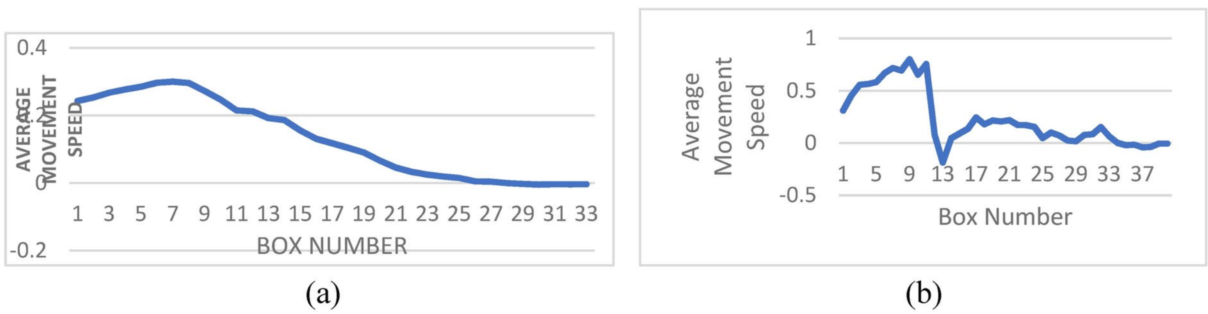


**Fig. 4**. Average Movement Speed between Box number and Average movement Speed for (**a**) Participant 1 and (**b**) Participant 2.

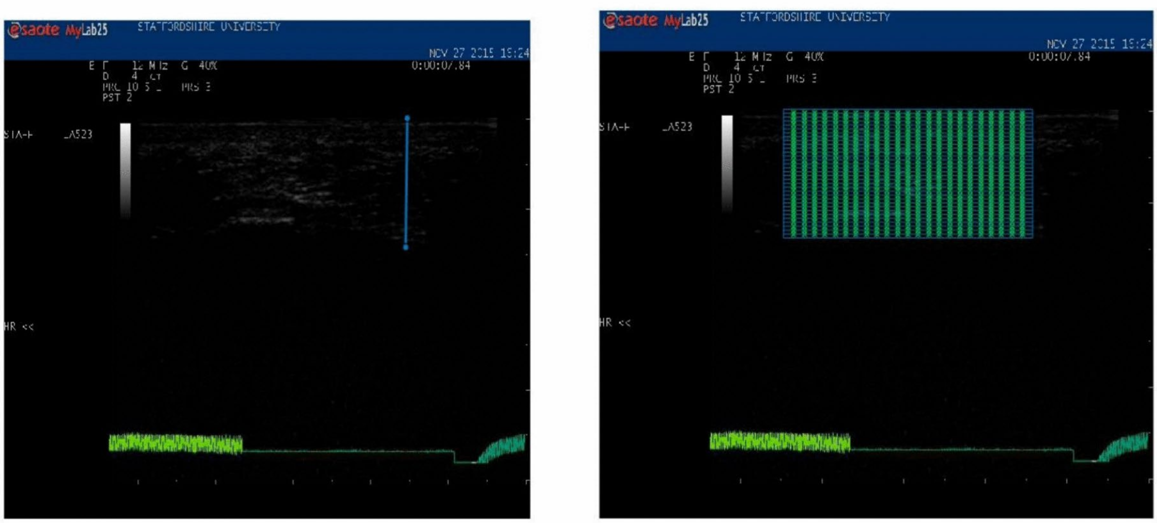

**Fig. 5**. Box segment Initialisation and automated box generation.

frame are ignored. The centroid values in a column correspond to the displacement values in successive video frames, and the next column refers to the next box in the box segment. Each box segment value is stored in separate Excel sheets. The number of Excel sheets available at the end of the tracking refers to the number of box segments in the image.

Then, in step 3, from the centroid values computed, the individual displacement of boxes over time is calculated in the first step. Then, the average value of displacement, also called average movement speed, is calculated. This is repeated for all the box segments in the image, and then the average movement speed is computed for each box segment.

The bone moves with the highest speed in the ultrasound video of the heel since the probe is pressed against the skin, and then the force is removed. The graph plot for average movement speed is calculated with the box number on the x-axis and average movement speed on the y-axis. This plot is calculated for all the box segments in the image.

When a box segment is drawn at the right corner of the image, boxes are automatically constructed along consecutive box segments across the entire image. Figure 7 corresponds to a box segment containing the bone region, represented by 33 boxes. The leftmost boxes (1–10) in the graph denote the bone region. The subsequent boxes (11 onward) represent the macro-chamber region, which eventually transitions to the micro-chamber region (boxes 29–33). The boxes in the stiffer micro-chamber region exhibit minimal movement, resulting in near-zero average movement speed. The macro-chamber region shows moderate movement, while the bone region, the fastest-moving structure relative to the probe, has the highest speed. A distinct difference is observed in Fig. 7 of the revised manuscript, where the highest peak indicates the bone region and zero movements indicate the micro-chamber layer. In contrast, this clear distinction is absent in Fig. 6, as it corresponds to a box segment where the bone region is absent.

Some of the box segments in the image do not have bone regions, as shown clearly in the plot (Fig. 6); those plots do not have a clear transition from the bone to non-bone areas. At the same time, the box segments with bone regions have a clear transition from bone to the non-bone areas (skin region -nearly zero speed and bone region – high speed.), as shown in Fig. 7. This concept is utilised to extract the bone from the surrounding tissues. The highest value from the average movement speed gives the bone region.

## Chan–Vese segmentation

Chan–Vese segmentation belongs to the model-based segmentation class, where the contour evolution is based on region information. Here, the initial unknown curve deforms and fits exactly on the region-of-interest thereby segmenting it. The process in Chan–Vese method is given by the Eq. (1).

$$\begin{aligned}\mathrm{GCV}\left(h_{1+}h_{2,}\vartheta\right) = G^{FT}\left(h_1+h_2,\vartheta\right)+\mu L\left(\vartheta\right)\left(\int_{\Omega}\left|S_m-h_1\right|^2 H\left(\vartheta\right)\mathrm{dx\ dy}\right.\\ +\int_{\Omega}\left|S_m-h_2\right|^2\left(1-H\left(\vartheta\right)\mathrm{dxdy}\right)+\mu\int_{\Omega}\left|\nabla_{\vartheta}\right|\ \mathrm{dx\ dy}\end{aligned} \tag{1}$$

where $G^{FT}$ represents the external energy of the curve C. $h_1, h_2$ are constants, and $\vartheta$ is a curve undefined, $H\left(\vartheta\right)$ is the Heaviside function, and $\nabla_{\vartheta}$ is the Dirac one-dimensional function. L and $\mu$ are the regularisation or penalty function associated with the level set function $\vartheta$ and range of fixed parameter less than 1. The input image is represented by $S_m$ and the particular image coordinate is given by $S_m\left(x,y\right).\nabla$ denotes the gradient operator and GCV $\left(h_{1+}h_{2,}\vartheta\right)$ represents the Chan–Vese function. The Chan–Vese method comes under the class of deformable models that deform under the influence of internal and external forces for delineating the region of interest. In the first step, a close curve is initialised within the region of interest. Internal forces maintain

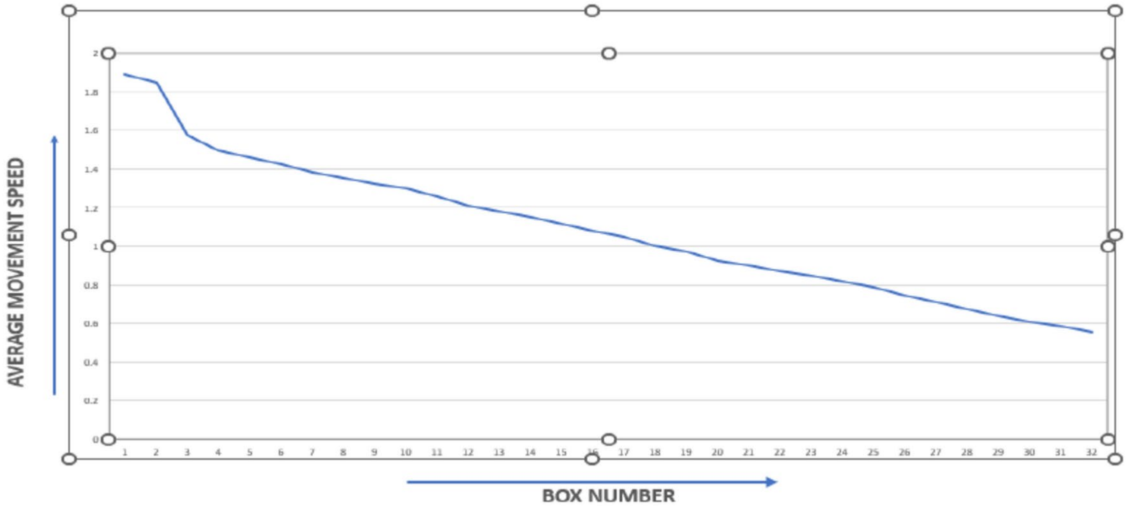


**Fig. 6**. Plot of average movement speed in non-bone regions.

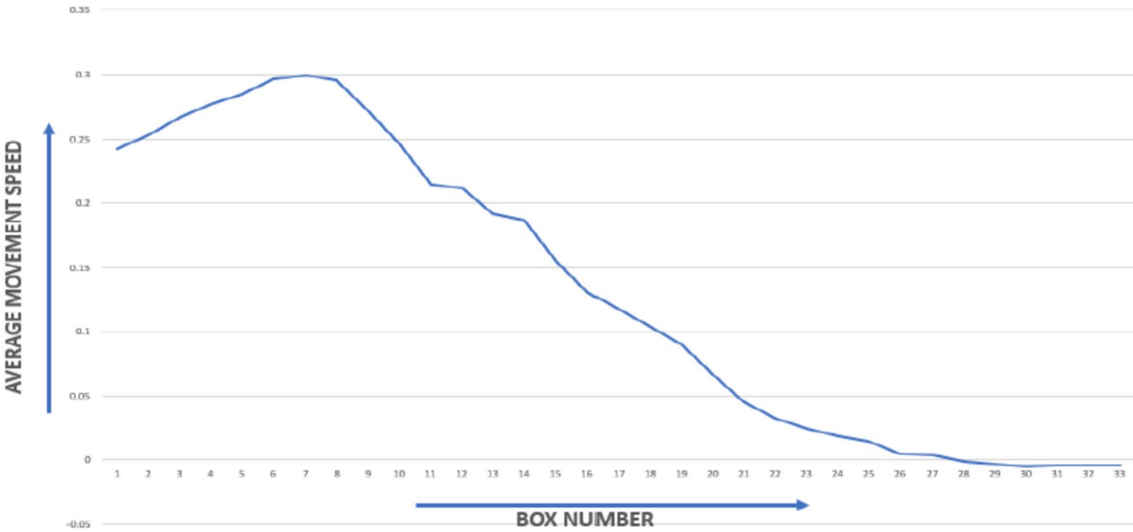


**Fig. 7**. Plot of average movement speed in bone regions.

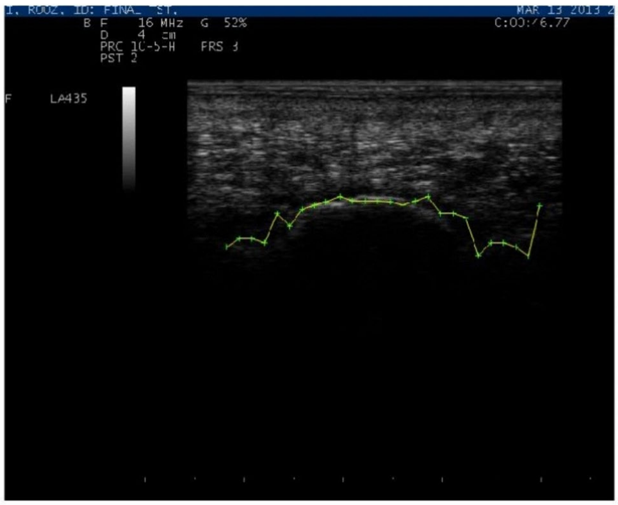

**Fig. 8**. Approximate bone region from the average movement speed.

the smoothness of the curve during deformation and are computed within the curve. External forces drive the curve towards the region to be segmented and are computed from the image. In edge-based deformable models (Snake models), the force generated for deformation towards the structure to be segmented is based on gradient information. In region-based deformable models (Chan–Vese method), force generated for deformation towards the structure to be segmented is based on region information. The advantage of the Chan–Vese method is less sensitive to noise and can segment the weaker boundaries in a better manner.

Chan–Vese method[78,79] is the extension of the traditional active contour algorithm[80], which comes under the class of model-based segmentation. The drawback of the active contour method is that the initial contour should be as close as possible to the region of interest, and the edge information leads to the detection of false edges in ultrasound imaging[42]. It is susceptible to noise. Chan–Vese segmentation uses region information for contour deformation and deforms the undefined curve even in noisy conditions. The initial contour obtained from the average movement speed values is segmented using the Chan–Vese algorithm, which is post-processed to get the final bone surface.

## Results

The average movement speed is calculated for all the box segments in the image. From the maximum value computed, find the maximum value from the computed maximum average speed value, termed 'a'. To identify the bone region from the box segments, it is essential to remove the non-bone regions. Hence, apply thresholding i.e., remove the speed values lying 25% of the 'a'. Compute the box coordinates corresponding to the maximum speed values and join them, as shown in Fig. 8. Then, traverse 30 pixels below each coordinate and join all the coordinates to form the closed curve, as shown in Fig. 9a. This curve is given as an initial contour for the Chan–Vese method. The segmented output from the Chan–Vese method is shown in Fig. 9b. Then, this output is post-processed with blob extraction (obtaining the largest area) and boundary tracing to get the final segmented bone surface, as shown in Fig. 9c. The accuracy of the segmented output from the algorithm is validated by comparison against the manually delineated bone image from the experts. Visual comparison between the algorithmic output and manually segmented output is shown 9d, which shows good agreement for bone boundary (yellow colour – manually segmented curve, green colour – algorithmic output).

Bhattacharyya distance was computed between algorithmic and manual segmentation to show the quantitative agreement between manual and algorithmic output. The Bhattacharyya distance between the manual and algorithmic output is 0.0643 for participant 1 (almost close to 0) and 0.0614 for participant 2, showing good quantitative agreement between the two outputs. Bhattacharyya distance measures the amount of overlap/similarity between two probability distributions. To this end, each contour is considered as a probability distribution. The Bhattacharyya distance is defined by Eq. (2). The closer the Bhattacharyya distance is to zero, the higher the level of agreement between contours.

$$\mathrm{DB}\ (\mathrm{l},\mathrm{m}) = \frac{1}{4}\mathbf{ln}\left(\frac{1}{4}\left(\frac{\sigma_l^2}{\sigma_m^2}+\frac{\sigma_m^2}{\sigma_l^2}+2\right)\right)+\frac{1}{4}\left(\frac{\left((\mu_l)-(\mu_m)\right)^2}{\sigma_l^2+\sigma_m^2}\right) \tag{2}$$

where $\sigma_l^2$, $\sigma_m^2$ is the variance of $l^{th}$ and $m^{th}$ distribution respectively, $\mu_{l,}$ $\mu_m$ the mean of the $l^{th}$ and $m^{th}$ distribution respectively, and l, m are two contour distributions.

In addition to the Bhattacharyya distance, the correlation coefficient between computerised and manual segmentation is computed. The correlation coefficient tells the linear relationship between two-line segments.

The following Eqs. (3–6) have estimated all three correlation parameters.

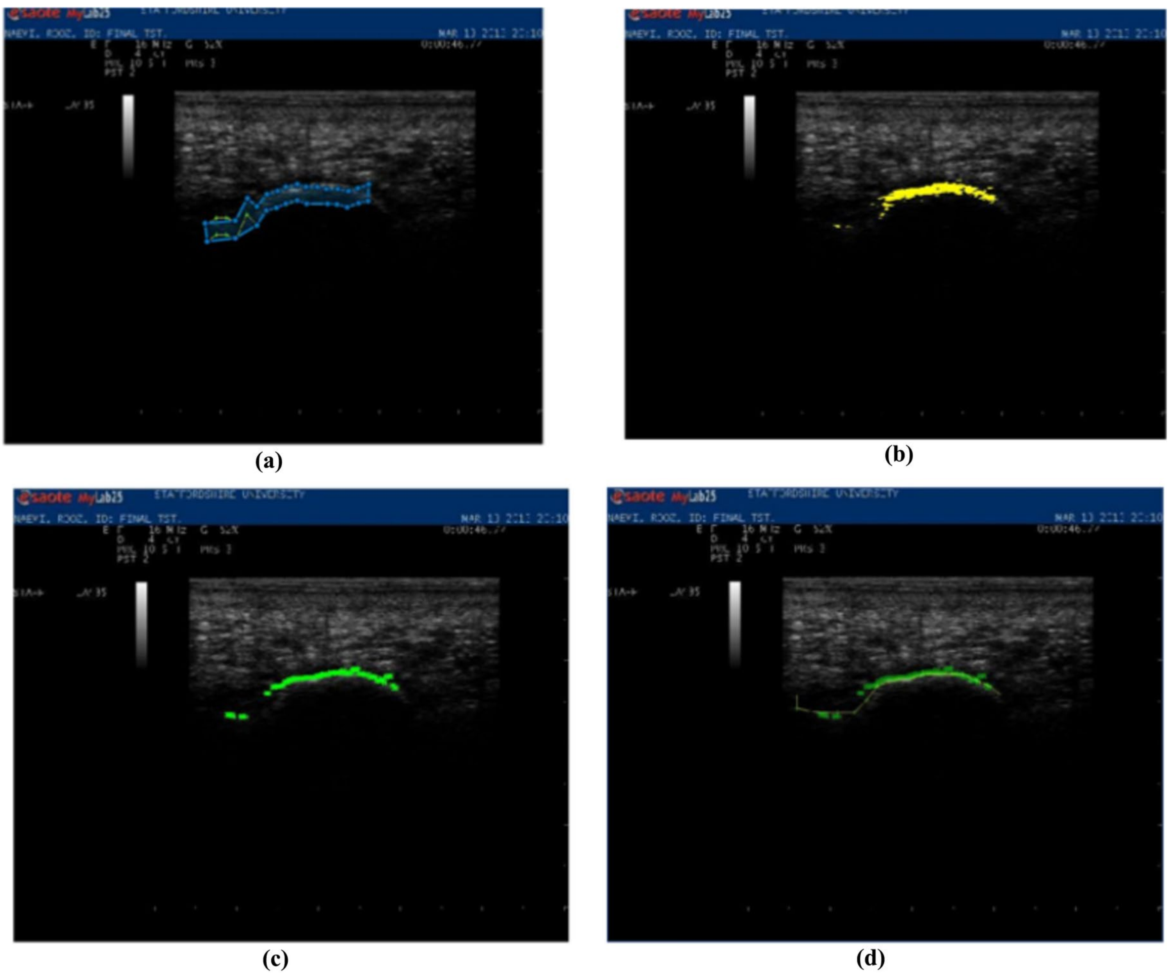


**Fig. 9**. Participant 1: (**a**) Contour initialisation in input image (**b**) output from Chan–Vese method (**c**) output after post-processing (**d**) overlaying of algorithmic output over manual segmentation for qualitative analysis.

$$E(x) = \frac{1}{N}\sum_{i=1}^{N} xi \tag{3}$$

$$D(x) = \frac{1}{N}\sum_{i=1}^{N} (xi - E(x))^2 \tag{4}$$

$$Cov\,(x, y) = \frac{1}{N}\sum_{i=1}^{N} (xi - E(x))(yi - E(y)) \tag{5}$$

$$\gamma xy = \frac{Cov\,(x, y)}{\sqrt{D(x)}\,\sqrt{D(y)}} \tag{6}$$

where x,y—Gray level values of two adjacent pixels, N—Total number of pixels,$\gamma xy -$ Cross correlation of two pixels.

The correlation coefficient value should be close to 1 for a good correlation between two-line segments. The computed value between manual and computerised segmentation is 0.9911 for Participant 1 and 0.9970 for Participant 2, which is close to 1, indicating a good correlation between them. The segmented output for Participant 2 is shown in Fig. 10.

The overall process for the bone segmentation algorithm is shown in Supplementary material C.

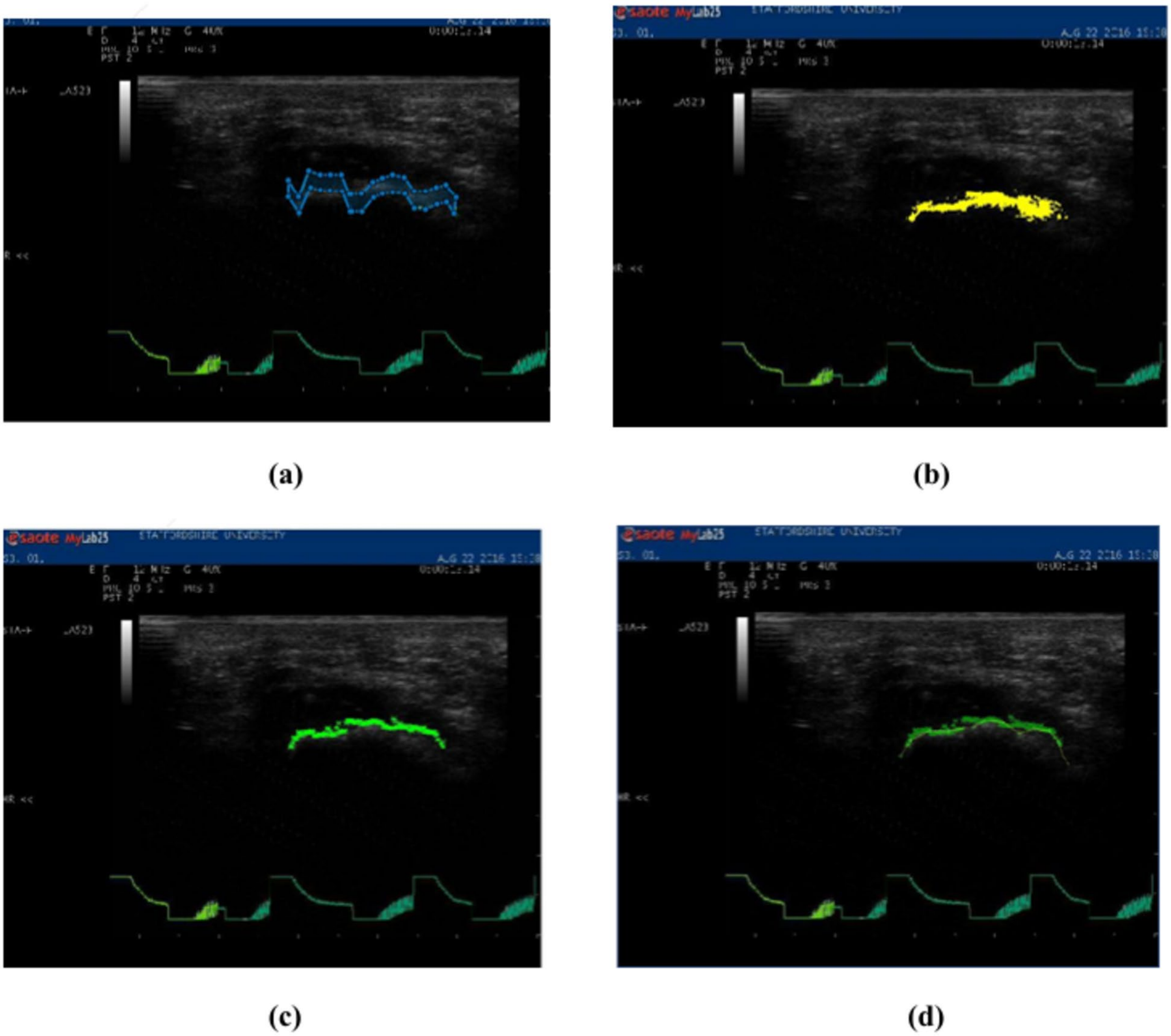


**Fig. 10**. Participant 2: (**a**) Contour initialisation in input image (**b**) output from Chan–Vese method (**c**) output after post-processing (**d**) overlaying of algorithmic output over manual segmentation for qualitative analysis.

In addition to the Bhattacharyya distance and correlation co-efficient, the Hausdorff distance has been chosen to prove the strength of the proposed algorithm.

The Hausdorff distance between two curves T and S is given by Eq. (7)

$$d_H\left(T,S\right) = max\left\{ {}_{t\in T}{}^{sup}{}_{s\epsilon S}{}^{inf} d\left(t,s\right), {}_{s\in S}{}^{sup}{}_{t\epsilon T}{}^{inf} d(t,s) \right\} \tag{7}$$

The Hausdorff distance calculates the largest minimum distance from points from one curve to the nearest point in another curve. The sets T & S contain 11 points each. The largest minimum Euclidean distance is calculated between any two points from one set to another. Upon conducting this experimentation for this proposed work, the Hausdorff distance was calculated as 11.5916 pixels for the first test sample and 24.0674 for the second video sample, respectively. Observing this shows that the largest minimum distance itself is significantly less, which ensures the efficacy of the fully automated segmentation process.

## Discussion

The optical flow algorithm is used to track the movement of interest points (inside the box) between two frames, and then the centroid value of the box is computed. Each box's average movement speed values are calculated from the computed centroid values. The box value with the highest average movement speed from the box segment corresponds to the bone region since bone moves faster than surrounding tissues in the ultrasound video. This work is the first attempt to use a tracking algorithm to segment the bone. While recording the ultrasound video of the bone, force is applied to the bone's heel region, enabling the bone's differentiation from the surrounding areas. The application of force while recording ultrasound videos of bones and using a tracking algorithm for bone segmentation is the novelty of this work.

The quantitative results between the algorithmic and manual segmentation are obtained using the Bhattacharyya distances and Correlation coefficient, respectively, for Participant 1 and Participant 2. The Bhattacharyya distance quantifies the degree of dissimilarity between two probability distributions. A value close to zero indicates that the two distributions overlap highly. The correlation coefficient quantifies the linear relationship between two data sets, with its value ranging from −1 to 1. A value close to one indicates that the two curves are well-correlated and represent a strong positive correlation. The Bhattacharyya distances for Participant 1 & Participant 2 are 0.0643 and 0.0614, respectively, almost close to 0. The correlation coefficient values for Participant 1 & Participant 2 are 0.9911 and 0.9970, respectively, almost close to 1. From the above results, it is evident that the computerised segmentation of bone for Participant 1 & Participant 2 has a high correlation with the manual segmentation, which shows the efficacy of the proposed algorithm. These results were compared with the segmentation results from the model[71] and texture[81] based methods, as shown in Table 1. The level set-based segmentation was used to segment the heelbone from the single frame of an ultrasound image. The level set-based methods come under the class of geometric deformable models, where a contour is initialised inside the image to be segmented. Based on the curve evolution, the initial contour fits precisely in the region of interest, thereby segmenting the heel bone. Unlike the traditional snake contour, which fails to handle the topological changes, the geometric deformable model handles the topological changes efficiently. To segment the heel bone, an initial contour is initialised to cover the heel bone region. The level set method fits the curve in the bone region based on curve evolution. The segmented outputs are compared with the manually segmented results from the experts, and accuracy is computed using Bhattacharyya distances.

The Bhattacharyya distances from the model-based segmentation are 1.0927 and 1.5233 for participant 1 and participant 2. The texture features are extracted from the heel bone using Gabor filters in texture-based segmentation. The extracted texture features are given to the K-means clustering algorithm to get the final segmented bone region. The segmented outputs are compared with the manually segmented results from the experts, and accuracy is computed using Bhattacharyya distances. The Bhattacharyya distances from the texture-based segmentation are 5.3961 and 0.5581 for participant 1 and participant 2. The closer the Bhattacharyya distances to zero, the higher the segmentation accuracy. The Bhattacharyya distances from the model and texture-based segmentation are very high compared to the proposed optical flow-based segmentation, as shown in Table 1. This indicates that the proposed optical flow-based segmentation has higher accuracy than the model and texture-based segmentation. The output results of texture and model-based segmentation methods for participant 1 and participant 2 are shown in Figs. 11 and 12.

To analyse the time complexity of the overall algorithm, the time complexity of automatic box construction, tracking of KAZE interest points and Chan–Vese segmentation algorithm are analysed. The automatic box construction takes the time complexity of O[KN], where K is the total number of frames tracked and N is the number of pixels in the frame. The tracking of KAZE interest points takes the time complexity of O[FN][82], where

| S. no. | Participant | Proposed optical flow-based segmentation accuracy | Segmentation accuracy from model-based segmentation | Segmentation accuracy from texture-based segmentation |
|---|---|---|---|---|
| | Participant 1 | 0.0643 | 1.0927 | 5.3961 |
| 2 | Participant 2 | 0.0614 | 1.5233 | 0.55807 |

**Table 1**. Comparison of proposed optical flow-based segmentation with model and texture-based segmentation.

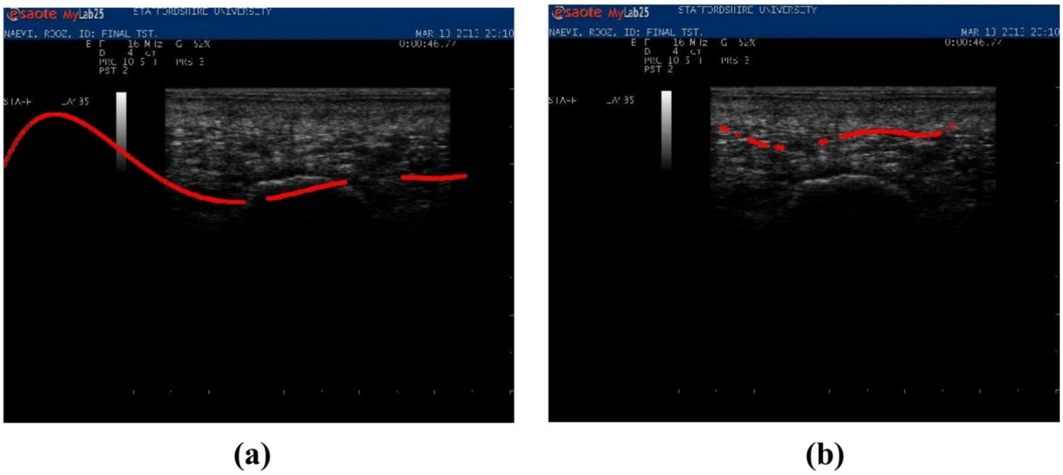


**Fig. 11**. (**a**) Model based segmentation of Participant 1 (**b**) Texture based segmentation of Participant 1.

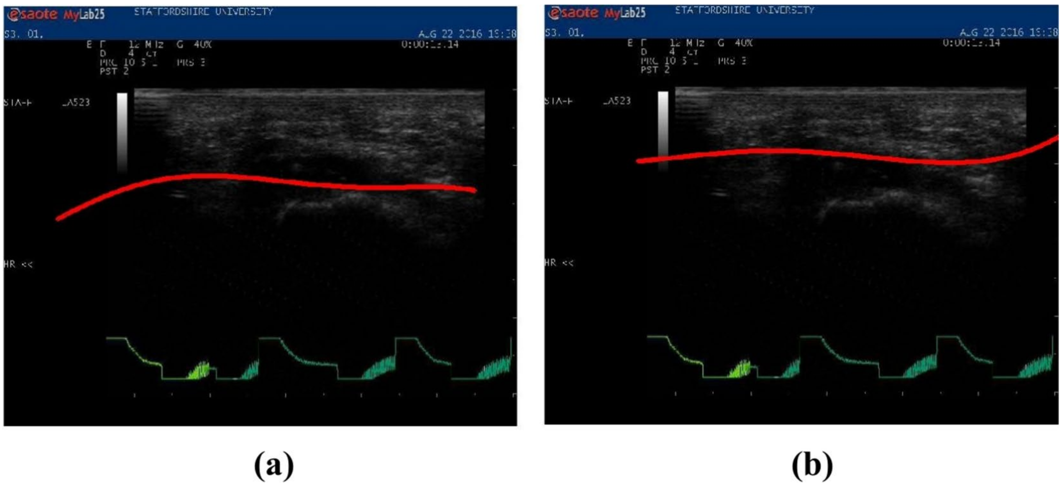


**Fig. 12**. (**a**) Model-based segmentation of Participant 2 (**b**) Texture-based segmentation of Participant 2.

F is the features and N is the quantity of pixels on the frame. The Chan–Vese segmentation algorithm takes a time complexity of O[N][83]. Thus, our proposed approach's total time complexity is O[(K + F)N].

This paper presents a concept which, for the first time, opens the way for the fully automated segmentation of bone based on movement. To this end, this work showcases that bones can be segmented based on motion information for the first time, complimenting model and texture-based segmentation. This proof-of-concept paper proposes a new algorithmic framework for movement-based bone segmentation. Due to limited resources, we could not analyse more cases. However, this study will be extended to include more videos in future research.

This proposed work has the use cases for the following targets:

- The development of the proposed method facilitates the automatic segmentation of the tendon from 2D ultrasound images. By interpolating a series of these 2D segmented images over time, a 3D geometry of the tendon can be reconstructed. This 3D geometry under varying loading conditions aids in analysing the tendon's regional adaptation to different exercise programs and design exercise rehabilitation strategies.
- The proposed algorithm can be explored in tissues exhibiting more complex movements and interactions with surrounding structures.
- Motion-based segmentation analyses frames of ultrasound images, and hence, it processes lot of images of cohorts, reducing the processing time to minutes. Therefore, this approach applies to clinics.
- Radiologists determine the tissue and it's boundary by analysing the frames of the video and by twisting & turning the probe of the ultrasound[84, 85]. This work aims to mimic the work of radiologists by using motion information for segmentation.

## Conclusion

This study presents an automated method for segmenting the heel bone from ultrasound images, marking the first attempt to achieve bone segmentation using motion information. The process begins with a single box segment drawn at the right edge of the image, followed by automated box construction based on defined sizes, overlaps, and thresholds on average movement speed. These parameters enable accurate segmentation by isolating bone regions while eliminating non-bone regions, though further research is needed to generalise these parameters for other videos. The Bhattacharyya distances for Participant 1 (0.0643) and Participant 2 (0.0614) demonstrate the method's precision. The proposed optical flow approach outperforms model-based (1.0927 and 1.5233 for Participants 1 and 2) and texture-based (5.3961 and 0.55807 for Participants 1 and 2) segmentation methods, achieving better accuracy with a time complexity of $O[(K+F)N]$. The segmentation leverages the displacement differences between the bone and the microchamber layer, arising from skin deformation and release, to differentiate bone from surrounding tissues. However, the current algorithm cannot detect the presence or absence of bone, as it relies on speed differentials. Future improvements should incorporate bone and shadow features to enhance the algorithm's ability to confirm bone presence in ultrasound videos.

## Data availability

The video used in this study can be found in Supplementary Material A. The X and Y coordinates of the tracked centroid pixels are given in Supplementary Material B. The overall flowchart for Heel bone segmentation is given in Supplementary Material C. The datasets analysed during the current study are available from the corresponding author on reasonable request.

## Code availability

Available upon request.

Received: 17 September 2024; Accepted: 27 February 2025
Published online: 08 March 2025

## References

1. Hacihaliloglu, I. Ultrasound imaging and segmentation of bone surfaces: A review. *Technology (Singap. World Sci.)* **5**, 74–80. https://doi.org/10.1142/S2339547817300049 (2017).
2. Peng, T. et al. Organ boundary delineation for automated diagnosis from multi-center using ultrasound images. *Expert Syst. Appl.* **238**, 122128. https://doi.org/10.1016/j.eswa.2023.122128 (2024).
3. Lyu, Y. et al. AMS-PAN: Breast ultrasound image segmentation model combining attention mechanism and multi-scale features. *Biomed. Signal Process. Control* **81**, 104425. https://doi.org/10.1016/j.bspc.2022.104425 (2023).
4. Migeon, B. In vitro3D reconstruction of long bones using B-scan image processing. *Med. Biol. Eng. Comput.* **35**, 369–372 (1997).
5. Behforootan, S., Chatzistergos, P. E., Chockalingam, N. & Naemi, R. A clinically applicable non-invasive method to quantitatively assess the visco-hyperelastic properties of human heel pad, implications for assessing the risk of mechanical trauma. *J. Mech. Behav. Biomed. Mater.* **68**, 287–295. https://doi.org/10.1016/j.jmbbm.2017.02.011 (2017).
6. Parashar, A., Rishi, R., Parashar, A. & Rida, I. Medical imaging in rheumatoid arthritis: A review on deep learning approach. *Open Life Sci.* https://doi.org/10.1515/biol-2022-0611 (2023).
7. Hemalatha, R. J., Vijaybaskar, V. & Thamizhvani, T. R. Automatic localisation of anatomical regions in medical ultrasound images of rheumatoid arthritis using deep learning. *Proc. Inst. Mech. Eng. Part H: J. Eng. Med.* **233**(6), 657–667. https://doi.org/10.1177/0954411919845747 (2019).
8. Fradi, M., Elhadj Youssef, W., Bouallegue, G., Machhout, M. & Lasaygues, P. Automatic USCT image processing segmentation for osteoporosis detection. In *Smart Innovation, Systems and Technologies*, Vol. 146. https://doi.org/10.1007/978-3-030-21005-2_36 (2020).
9. Laugier, P., Lefebvre, F., Roux, C. & Berger, G. Segmentation of QUS images of the Calcaneus using elastic deformation of flexible Fourier contour. In *IEEE Ultrasonics Symposium*-1225, 1225–1228 (2001).
10. Lopez-Perez, L., Lemaitre, J., Alfiansyah, A. & Bellemare, M.-E. Bone surface reconstruction using localised freehand ultrasound imaging. In *Conference of the IEEE Engineering in Medicine and Biology Society*, Vol. 2008, 2964–2967. https://doi.org/10.1109/IEMBS.2008.4649825 (2008).
11. Veronese, E. et al. Improved detection of synovial boundaries in ultrasound examination by using a cascade of active-contours. *Med. Eng. Phys.* **35**, 188–194. https://doi.org/10.1016/j.medengphy.2012.04.014 (2013).
12. Martins, N., Saad Sultan, M. Veiga, D. et al. Segmentation of the metacarpus and phalange in musculoskeletal ultrasound images using local active contours. In *Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 4097–4100. https://doi.org/10.1109/EMBC.2016.7591627 (2016).
13. Luis-garcia, R. de, & Alberola-Lopez, C. Parametric 3D hip joint segmentation for the diagnosis of developmental dysplasia, 4807–4810 (2006).
14. Tran, D. & Rohling, R. N. Automatic detection of lumbar anatomy in ultrasound images of human subjects. *IEEE Trans. Biomed. Eng.* **57**, 2248–2256. https://doi.org/10.1109/TBME.2010.2048709 (2010).
15. Hacihaliloglu, I., Abugharbieh, R., Hodgson, A. & Rohling, R. Bone segmentation and fracture detection in ultrasound using 3D local phase features bone segmentation and fracture detection in ultrasound using 3D local phase features. In *Medical Image Computing and Computer-Assisted Intervention*, 287–295 (2008).
16. Pandey, P., Guy, P., Hodgson, A. J. & Abugharbieh, R. Fast and automatic bone segmentation and registration of 3D ultrasound to CT for the full pelvic anatomy: a comparative study. *Int. J. Comput. Assist. Radiol. Surg.* **13**, 1515–1524. https://doi.org/10.1007/s11548-018-1788-5 (2018).
17. Hacihaliloglu, I., Guy, P., Hodgson, A. J. & Abugharbieh, R. Volume-specific parameter optimisation of 3D local phase features for improved extraction of bone surfaces in ultrasound. *Int. J. Med. Robot. Comput. Assist. Surg.* **10**, 461–473. https://doi.org/10.1002/rcs.1552 (2014).
18. Sultan, M. S., Martins, N., Ferreira, M. J. & Coimbra, M. T. Segmentation of bones & MCP joint region of the hand from ultrasound images. In *IEEE Engineering in Medicine and Biology Society Conference*, 3001–3004 (2015).
19. Baka, N., Leenstra, S. & van Walsum, T. Random forest-based bone segmentation in ultrasound. *Ultrasound. Med. Biol.* **43**, 1–12. https://doi.org/10.1016/j.ultrasmedbio.2017.04.022 (2017).
20. Faisal, A., Ng, S. C., Goh, S. L. & Lai, K. W. Knee cartilage segmentation and thickness computation from ultrasound images. *Med. Biol. Eng. Comput.* **56**, 657–669. https://doi.org/10.1007/s11517-017-1710-2 (2018).

21. Villa, M. et al. FCN-based approach for the automatic segmentation of bone surfaces in ultrasound images. *Int. J. Comput. Assist. Radiol. Surg.* **13**, 1707–1716. https://doi.org/10.1007/s11548-018-1856-x (2018).
22. Alsinan, A. Z., Patel, V. M. & Hacihaliloglu, I. Automatic segmentation of bone surfaces from ultrasound using a filter-layer-guided CNN. *Int. J. Comput. Assist. Radiol. Surg.* **14**, 775–783. https://doi.org/10.1007/s11548-019-01934-0 (2019).
23. Wang, P., Wang, P., Vives, M., Patel, V. M. & Hacihaliloglu, I. Robust real-time bone surfaces segmentation from ultrasound using a local phase tensor-guided CNN. *Int. J. Comput. Assist. Radiol. Surg.* **15**, 1127 (2020).
24. Hemalatha, R. J., Vijaybaskar, V. & Thamizhvani, T. R. Automatic localisation of anatomical regions in medical ultrasound images of rheumatoid arthritis using deep learning. *Proc. Inst. Mech. Eng. Part H: J. Eng. Med.* **233**, 657–667. https://doi.org/10.1177/0954411919845747 (2019).
25. Luan, K. & Zeyu Li, J. L. An efficient end-to-end CNN segmentation of bone surfaces from ultraasound. *Comput. Med. Imaging Graph.* **84**, 1–8 (2020).
26. Xu, M., Ma, Q., Zhang, H., Kong, D. & Zeng, T. MEF-UNet: An end-to-end ultrasound image segmentation algorithm based on multi-scale feature extraction and fusion. *Comput. Med. Imaging Graph.* **114**, 102370. https://doi.org/10.1016/j.compmedimag.2024.102370 (2024).
27. Lyu, J. et al. Dual-task ultrasound spine transverse vertebrae segmentation network with contour regularisation. *Comput. Med. Imaging Graph.* **89**, 101896. https://doi.org/10.1016/j.compmedimag.2021.101896 (2021).
28. Ramakrishnan, A. H. et al. A concept for movement-based computerised segmentation of connective tissue in ultrasound imaging. *Multimed. Tools Appl.* https://doi.org/10.1007/s11042-022-12932-x (2022).
29. Yang, X., Tang, S., Tasciotti, E. & Righetti, R. Bone surface enhancement in ultrasound images using a new Doppler-based acquisition/processing method. *Phys. Med. Biol.* **63**(2), 025035. https://doi.org/10.1088/1361-6560/aa9c5e (2018).
30. Ramakrishnan, A. H., Rajappa, M., Kirthivasan, K., Chockalingam, N., Chatzistergos, P. E. & Amirtharajan, R. A systematic survey on segmentation algorithms for musculoskeletal tissues in ultrasound imaging. In *Archives of Computational Methods in Engineering*. Springer Science and Business Media B.V. https://doi.org/10.1007/s11831-024-10171-x (2024).
31. Kass, M., Witkin, A. & Terzopoulos, D. Snakes: Active contour models. *Int. J. Comput. Vis.* **1**, 321–331. https://doi.org/10.1007/BF00133570 (1988).
32. Berton, F., Cheriet, F., Miron, M. C. & Laporte, C. Segmentation of the spinous process and its acoustic shadow in vertebral ultrasound images. *Comput. Biol. Med.* **72**, 201–211. https://doi.org/10.1016/j.compbiomed.2016.03.018 (2016).
33. Wang, X. F., Min, H., Zou, L. & Zhang, Y. G. A novel level set method for image segmentation by incorporating local statistical analysis and global similarity measurement. *Pattern Recognit.* **48**(1), 189–204. https://doi.org/10.1016/j.patcog.2014.07.008 (2015).
34. Guo, Y., Du, G. Q., Xue, J. Y., Xia, R. & Wang, Yh. A novel myocardium segmentation approach based on neutrosophic active contour model. *Comput. Methods Programs Biomed.* **142**, 109–116. https://doi.org/10.1016/j.cmpb.2017.02.020 (2017).
35. Tsechpenakis, G. *Deformable Model-Based Medical Image Segmentation* (2011).
36. Hesamian, M. H., Jia, W., He, X. & Kennedy, P. Deep learning techniques for medical image segmentation: Achievements and challenges. *J. Dig. Imaging* **32**, 582–596. https://doi.org/10.1007/s10278-019-00227-x (2019).
37. Jiang, W. et al. Ori-Net: Orientation-guided neural network for automated coronary arteries segmentation. *Expert Syst. Appl.* **238**, 121905. https://doi.org/10.1016/j.eswa.2023.121905 (2024).
38. Liu, S. et al. Deep learning in medical ultrasound analysis: A review. *Engineering* **5**, 261–275. https://doi.org/10.1016/j.eng.2018.11.020 (2019).
39. Mansoor, A. et al. Segmentation and image analysis of abnormal lungs at CT: Current approaches, challenges, and future trends. *Radiographics* **35**, 1056–1076. https://doi.org/10.1148/rg.2015140232 (2015).
40. Lucas, B. D. & Kanade, T. An iterative image registration technique with an application to stereo vision. In *Proceedings of the 7th International Joint Conference on Artificial Intelligence (IJCAI '81)*, 121–130 (1981).
41. Laurence, N. & Bohs, G. E. T. A novel method for angle independent ultrasonic imaging of blood flow and tissue motion. *IEEE Trans. Biomed. Eng.* **38**, 280–286 (1991).
42. Paramkusam, A. V. & Reddy, V. S. K. An efficient fast full search block matching algorithm with SSD criterion. In *2011 Annual IEEE India Conference: Engineering Sustainable Solutions, INDICON-2011*, 3–8 (IEEE, 2011).
43. Golemati, S. et al. Carotid artery wall motion estimated from B-mode ultrasound using region tracking and block matching. *Ultras. Med. Biol.* **29**, 387–399. https://doi.org/10.1016/S0301-5629(02)00760-3 (2003).
44. Chernak, L. A. & Thelen, D. G. Tendon motion and strain patterns evaluated with two-dimensional ultrasound elastography. *J. Biomech.* **45**, 2618–2623. https://doi.org/10.1016/j.jbiomech.2012.08.001 (2012).
45. Drazan, J. F., Hullfish, T. J. & Baxter, J. R. An automatic fascicle tracking algorithm quantifying gastrocnemius architecture during maximal effort contractions. *PeerJ* https://doi.org/10.7717/peerj.7120 (2019).
46. Bharadwaj, S. & Almekkawy, M. Deep learning based motion tracking of ultrasound image sequences. In *IEEE International Ultrasonics Symposium, IUS*, 2020-September. https://doi.org/10.1109/IUS46767.2020.9251739 (2020).
47. Evain, E. et al. A pilot study on convolutional neural networks for motion estimation from ultrasound images. *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **67**(12), 2565–2573. https://doi.org/10.1109/TUFFC.2020.2976809 (2020).
48. Liu, F. et al. Cascaded one-shot deformable convolutional neural networks: Developing a deep learning model for respiratory motion estimation in ultrasound sequences. *Med. Image Anal.* **65**, 101793. https://doi.org/10.1016/j.media.2020.101793 (2020).
49. Wei, X. et al. Unsupervised convolutional neural network for motion estimation in ultrasound elastography. *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **69**(7), 2236. https://doi.org/10.1109/TUFFC.2022.3171676 (2022).
50. De Luca, V., Székely, G. & Tanner, C. Estimation of large-scale organ motion in B-mode ultrasound image sequences: A survey. *Ultras. Med. Biol.* **41**, 3044–3062. https://doi.org/10.1016/j.ultrasmedbio.2015.07.022 (2015).
51. Korstanje, J. W. H. et al. Development and validation of ultrasound speckle tracking to quantify tendon displacement. *J. Biomech.* **43**, 1373–1379. https://doi.org/10.1016/j.jbiomech.2010.01.001 (2010).
52. Ophir, J., Céspedes, I., Ponnekanti, H., Yazdi, Y. & Li, X. Elastography: A quantitative method for imaging the elasticity of biological tissues. *Ultrason. Imaging* **13**(2), 111. https://doi.org/10.1177/016173469101300201 (1991).
53. Revell, J., Mirmehdi, M. & McNally, D. Computer vision elastography: Speckle adaptive motion estimation for elastography using ultrasound sequences. *IEEE Trans. Med. Imaging* **24**(6), 755. https://doi.org/10.1109/TMI.2005.848331 (2005).
54. Chuang, B. I. et al. Tendon-motion tracking in an ultrasound image sequence using optical-flow-based block matching. *BioMed. Eng. Online* **16**, 47. https://doi.org/10.1186/s12938-017-0335-x (2017).
55. Ayvali, E. & Desai, J. P. Optical flow-based tracking of needles and needle-tip localization using circular hough transform in ultrasound images. *Ann. Biomed. Eng.* **43**, 1828–1840. https://doi.org/10.1007/s10439-014-1208-0 (2015).
56. Tenbrinck, D. et al. Histogram-based optical flow for motion estimation in ultrasound imaging. *J. Math. Imaging Vis.* **47**, 138–150. https://doi.org/10.1007/s10851-012-0398-z (2013).
57. Li, Q. et al. Continuous detection of muscle aspect ratio using keypoint tracking in ultrasonography. *IEEE Trans. Biomed. Eng.* **60**, 2361–2369. https://doi.org/10.1109/TBME.2013.2256786 (2013).
58. Alvarez, L., Weickert, J. & Anchez, J. S.´ Reliable Estimation of Dense Optical Flow Fields with Large Displacements (2000).
59. Bouguet J-Y (1999) Pyramidal Implementation of the Lucas Kanade Feature Tracker Description of the algorithm
60. Delaunay, R., Hu, Y. & Vercauteren, T. An unsupervised learning approach to ultrasound strain elastography with spatio-temporal consistency. *Phys. Med. Biol.* **66**(17), 175031. https://doi.org/10.1088/1361-6560/ac176a (2021).

61. K. Z. Tehrani, A., Mirzaei, M. & Rivaz, H. Semi-supervised training of optical flow convolutional neural networks in ultrasound elastography. In *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 12263 LNCS. https://doi.org/10.1007/978-3-030-59716-0_48 (2020).
62. Tomasi, C. & Kanade, T. Detection and tracking of point features. School of Computer Science, Carnegie Mellon Univ 91, 1–22 (1991).
63. Wang, C., Zhang, J. & Liu, S. Medical ultrasound image segmentation with deep learning models. *IEEE Access* **11**, 10158–10168. https://doi.org/10.1109/ACCESS.2022.3225101 (2023).
64. Pandey, P. U., Hohlmann, B., Broessner, P., Hacihaliloglu, I., Barr, K., Ungi, T., Zettinig, O., Prevost, R., Dardenne, G., Fanti, Z., Wein, W., Stindel, E., Cosio, F. A., Guy, P., Fichtinger, G., Radermacher, K. & Hodgson, A. J. Standardised evaluation of current ultrasound bone segmentation algorithms on multiple datasets. In *Health Sciences EPiC Series in Health Sciences*, Vol. 5 (2022).
65. He, Q., Yang, Q. & Xie, M. HCTNet: A hybrid CNN-transformer network for breast ultrasound image segmentation. *Comput. Biol. Med.* **155**, 106629. https://doi.org/10.1016/j.compbiomed.2023.106629 (2023).
66. Chen, G., Li, L., Zhang, J. & Dai, Y. Rethinking the unpretentious U-net for medical ultrasound image segmentation. *Pattern Recognit.* **142**, 109728. https://doi.org/10.1016/j.patcog.2023.109728 (2023).
67. Zheng, T. et al. Segmentation of thyroid glands and nodules in ultrasound images using the improved U-Net architecture. *BMC Med. Imaging* https://doi.org/10.1186/s12880-023-01011-8 (2023).
68. Shereena, V. B. & Raju, G. Medical ultrasound image segmentation using multi-residual U-Net architecture. *Multimed. Tools Appl.* https://doi.org/10.1007/s11042-023-16461-z (2023).
69. Scott, I. et al. An automated method for tendon image segmentation on ultrasound using grey-level co-occurrence matrix features and hidden Gaussian Markov random fields. *Comput. Biol. Med.* **169**, 107872. https://doi.org/10.1016/j.compbiomed.2023.107872 (2024).
70. Ramakrishnan, A. H., Rajappa, M., Kannan, K., Narayanan, P. V. L., Chatzistergos, P. E. & Chockalingam, N. Segmentation of Patellar tendon from ultrasound images using active shape models. In *2019 10th International Conference on Computing, Communication and Networking Technologies, ICCCNT 2019*, 1–4. https://doi.org/10.1109/ICCCNT45670.2019.8944550 (2019).
71. Shu, X., Yang, Y., Liu, J., Chang, X. & Boying, W. ALVLS: Adaptive local variances-Based levelset framework for medical images segmentation. *Pattern Recognit.* **136**, 109257. https://doi.org/10.1016/j.patcog.2022.109257 (2023).
72. Rueda, S., Knight, C. L., Papageorghiou, A. T. & Alison Noble, J. Feature-based fuzzy connectedness segmentation of ultrasound images with an object completion step. *Med. Image Anal.* **26**(1), 30. https://doi.org/10.1016/j.media.2015.07.002 (2015).
73. Ramakrishnan, A. H. et al. A concept for movement-based computerised segmentation of connective tissue in ultrasound imaging. *Multimed. Tools Appl.* **81**(26), 38053–38066. https://doi.org/10.1007/s11042-022-12932-x (2022).
74. Hsu, C. C. et al. Microchambers and macrochambers in heel pads: Are they functionally different?. *J. Appl. Physiol.* **102**(6), 2227–2231. https://doi.org/10.1152/japplphysiol.01137.2006 (2007).
75. Naemi, R., Behforootan, S., Chatzistergos, P. & Chockalingam, N. Viscoelasticity in foot-ground interaction. In *Viscoelastic and Viscoplastic Materials*. InTech. https://doi.org/10.5772/64170 (2016).
76. Chatzistergos, P. E., Naemi, R., Sundar, L., Ramachandran, A. & Chockalingam, N. The relationship between the mechanical properties of heel-pad and common clinical measures associated with foot ulcers in patients with diabetes. *J. Diabetes Complic.* **28**(4), 488–493. https://doi.org/10.1016/j.jdiacomp.2014.03.011 (2014).
77. Behforootan, S., Chatzistergos, P. E., Chockalingam, N. & Naemi, R. A simulation of the viscoelastic behaviour of heel pad during weight-bearing activities of daily living. *Ann. Biomed. Eng.* **45**(12), 2750. https://doi.org/10.1007/s10439-017-1918-1 (2017).
78. Chan, T. F. & Vese, L. A. Active contours without edges. *IEEE Trans. Image Process.* **10**, 266–277 (2001).
79. Ramu, S. M. et al. A method to improve the computational efficiency of the Chan–Vese model for the segmentation of ultrasound images. *Biomed. Signal Process. Control* **67**, 102560. https://doi.org/10.1016/j.bspc.2021.102560 (2021).
80. Kass, M. & Witkin, A. *Snakes: Active Contour Models* (KIuwer Academic Publishers, 1988).
81. Li, H. A. et al. Facial image segmentation based on gabor filter. *Mathe. Problems Eng.* https://doi.org/10.1155/2021/6620742 (2021).
82. Baker, S. & Matthews, I. (n.d.). Lucas-Kanade 20 Years On: A Unifying Framework: Part 1.
83. Li, S. & Zhang, Q. Fast image segmentation based on efficient implementation of the Chan–Vese model with discrete gray level sets. *Math. Problems Eng.* https://doi.org/10.1155/2013/508543 (2013).
84. Blankstein, A. Ultrasound in the diagnosis of clinical orthopedics: The orthopedic stethoscope. *World J. Orthoped.* **2**, 13–24 (2011).
85. Bahner, D. P. et al. Language of transducer manipulation: Codifying terms for effective teaching. *J. Ultras. Med.* **35**, 183–188. https://doi.org/10.7863/ultra.15.02036 (2016).

## Acknowledgements
This work has been supported by the UK-India Education and Research Initiative (UKIERI) grant 'Ultrasound based assessment of tissue biomechanics to enhance the clinical management of foot related pathologies' (Project reference number: DST/INT/UK/P-145/2016). Financial support was also obtained from Department of Science and Technology (DST), New Delhi. The authors would also like to thank TATA Realty-IT City—SASTRA Srinivasa Ramanujam Research cell of SASTRA University

## Author contributions
Ananth Hari Ramakrishnan: Conceptualisation, Methodology, Software, Data curation, Writing- Original draft preparation. Muthaiah Rajappa, Nachiappan Chockalingam, Panagiotis E Chatzistergos, Krithivasan Kannan: Validation, Visualisation, Investigation, Data curation, Supervision. Rengarajan Amirtharajan: Validation, Writing- Reviewing and Editing. All authors reviewed the manuscript.

## Declarations

## Competing interests
The authors declare no competing interests.

## Ethical approval
All methods were carried out in accordance with relevant guidelines and regulations. Ethical approval was secured from Staffordshire University Ethics committee, Stoke-on-Trent, England. All participants provided written informed consent before any data was collected.

## Additional information
**Supplementary Information** The online version contains supplementary material available at https://doi.org/1

0.1038/s41598-025-92380-3.

**Correspondence** and requests for materials should be addressed to R.A.

**Reprints and permissions information** is available at www.nature.com/reprints.

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.