Literature review on wearable/mountable robots for positioning ultrasound (US) scanner

Rayan Hassan

University of Pittsburgh

Summer Undergraduate Research Program (SURP) – 2022

Supervisor: Dr. Hongliang Ren

Table of Contents

1. Abstract

Purpose of Review
This review provides information on recent robotic systems and applications used in ultrasound scanning. Following Lim et al.'s project on Collapsible Mount for Transcranial Doppler (COMPTRAD), it also presents AI methods used to localize targeted skin areas for US scanning. Although this paper isn't exactly a continuation of the COMPTRAD report, it offers thorough explanations that could help build upon any project involving the design and fabrication of US robotic devices. It touches on several aspects of US scanning and gives clear reviews on basic concepts used for in medical procedures.

Content
This paper starts with a detailed overview of ultrasound principles. Particularly, it tackles notions like sound wave propagation in different environments including soft tissues, wave interactions with surfaces like reflection and refraction, image resolution, and finally inner-workings of transducers. After that, it digs into robotic systems used for US scanning. Starting with a brief summary of probe positioning concepts, it mentions mountable, wearable and soft robots applied for ultrasound purposes. Their applications are later illustrated, ranging from human assisted to autonomous systems. Finally, a section on targeted areas' localization is presented. Starting with a review of Convolutional Neural Networks (CNNs), shows recent designs for AI algorithms that help target specific areas for several applications.

Keywords
Ultrasound scanning, probe positioning, mountable robots, wearable robots, soft robotics, origami device, teleoperation, robotic arm, autonomous robotic systems, CNN, convolutional layer, pooling layers, Fully Connected Layers (FCL), localized US scan

2. Introduction

Ultrasound (US) is a method that has been used throughout the years for biomedical purposes. It has been proved to be one of the safest, non-invasive imaging methods that has ever been invented. Unlike X-rays, MRIs and other imaging tools, US prevents the exposure to radiations and is more convenient and less costly. US corresponds to sound waves with high frequencies between 2 and 15 MHz, not detectable by the human ear. It works by sending out mechanical waves that will be transmitted through soft tissues and reflected on tissue surfaces to the transducer. The latter is held by a probe positioned on the patient's body in a way that optimizes image quality. A robotic system carries the probe and is connected to a screen that shows the radiologist US images collected in the moment. Although robotic arms are widely used for that purpose, wearable robots have also been explored, especially for brain scans. Lim et al. developed a COMTRAD device that registers Cerebral Blood Flow Velocity (CBFV) for early ischemic stroke detection [1]. It holds a Transcranial Doppler (TCD) ultrasound probe. The COMTRAD consists of a system that is placed onto the temporal window of the patient's head. It includes the TCD probe which is attached by 8 strings, 4 of which are attached to the top plate of the probe holder and 4 attached to its bottom plate. It contains straps that hold all the wires and string sheaths connecting to an external unit that has the microcontroller, 8 rotary motors and a battery away from the patient's head. The actuators allow for 4 Degrees of Freedom (DOF) ensured by the probe holder which exerts rotation about the roll and pitch axes. Another 1 DOF is secured along the vertical axis perpendicular to the skin plane. The results were successful and the probe was able to cover 7 cm^2 at a speed of approximatively 3.14 cm/s. This wearable device is an example of robotics employed for US scanning. Generally speaking, these robots have to showcase flexibility, accuracy and safety. Soft robots are used for the purpose of compliance as they are made off of soft, malleable materials which a large variety or achievable shapes. As for accuracy, it mostly involves the optimization of wave emission and localization of exact areas of interest. This enables sonographers to obtain higher quality images and facilitate drug deliveries through ultrasound waves. This paper is an informative review on what has already been accomplished in multiple aspects of US procedures and is divided into four main parts. Section I explains the physics of ultrasound emissions, section II cover different types of robots implemented in US scanning, section III tackles collaborative and autonomous robotic systems, and finally section IV deals with localized US scans.

3. Review on US principles

3.1 Propagation of Sound Waves

To understand how US scanning works, one must comprehend the behaviour of a sound wave in different materials. The propagation of these waves are dependent on several factors, like density and compressibility. The denser the material, the less sound propagates quickly. Similarly, the more compressible a material is, the slower sound propagates. The following can be represented by these formulas: $c \propto \frac{1}{\sqrt{\rho}}$; $c \propto \frac{1}{\sqrt{\kappa}}$. c is the acoustic velocity in m/s, ρ is the density in Kg/m^3 and κ is the compressibility [2]. The reciprocal of compressibility is stiffness, also called bulk modulus denoted by B:

$c \propto \frac{1}{\sqrt{\kappa\rho}}$; $c \propto \frac{\sqrt{B}}{\sqrt{\rho}}$ [2]. The stiffer a material is the faster sound propagates. For example, water is 13.4 times more compressible than mercury. In contrast, mercury is 13.6 times denser than water. Soft tissues on the other hand behave such that density and compressibility cancel each other out and velocities are similar for the most part (around 1540m/s). Generally, temperature does affect velocity but since it is constant across the body, it won't be a contributing factor.

3.2 Sound Waves Interactions

Sound waves interact with tissues in several ways. The one that counts most towards getting ultrasound images is reflection, where sound waves are reflected after hitting a surface. Each surface tissue has a density which allows us to differentiate between different tissues. Therefore, each wave has a unique acoustic impedance (Z) upon hitting a surface tissue, given by the following formula: $Z = \rho c$. The velocity being constant across the body as mentioned previously, the acoustic impedance of most tissues is primarily a function of the tissue density. An acoustic mismatch is created whenever waves hit adjacent regions with different densities and sound waves are reflected by the mismatch [3]. The greater the latter, the more echoes are reflected back. Consequently, ultrasound imaging is limited by the presence of bones and gas-filled structures. In fact, as sound is quickly transmitted through bones and poorly transmitted through air-filled structures, a high mismatch is created at bone and gas interfaces. Therefore, the majority of sound waves are reflected by to the transducer. This limits the penetration of sound waves into deeper tissues and causes image artefacts. The percentage reflectivity of waves is calculated using the following formula:

$$\%R = \left(\frac{Z_2 - Z_1}{Z_2 + Z_1}\right)^2 \times 100$$

The percent transmission is thus calculated by $\%T = 100 - \%R$

For example, $\%R = 43\%$ and $\%T = 57\%$ at the soft tissue – bone interface. Whereas $\%R = 99.9\%$ at the air – soft tissue interface, which explains the application of gel beforehand in order to eliminate air bubbles that prevent sound waves to be transmitted. Reflection works best on flat surfaces, where the entirety of the sound wave gets reflected back. However, as flat surfaces are hard to find within soft tissues, it is expected that fewer echoes are reflected back to the transducer because there are many different angles of reflection.

Another important notion is refraction, where a certain percentage of reflected sound gets through the tissue and propagates towards other tissues [3]. This causes multiple challenges in ultrasound scanning, one of which is the detection of unwanted or inexact matters and/or substances on the retrieved images.

Other equally important mechanisms occur in US like diffraction, interference and attenuation. However, for the purpose of simplicity, we won't be getting into each one of them in details.

3.3 Image Resolution

Resolution is the ability to accurately distinguish two closely situated structures. Axial resolution is being able to accurately distinguish between two close objects along the beam axis. Higher frequency means better axial resolution [2]. However, there is a trade-off that has to be made since depth decreases with increasing frequency because the signal gets attenuated. On the other hand, lateral resolution is the ability to differentiate between two objects adjacent to one another that are perpendicular to the beam axis. The narrower the beam width, the better the lateral resolution. Fig.1 illustrates how lateral and axial resolutions can be depicted in a US image. Spatial resolution is being able to display two objects close to one another as two separate images [4]. Finally, temporal resolution is the ability to distinguish two events in time [3]. It is determined by the number of image frames that can be acquired per second (Hz).

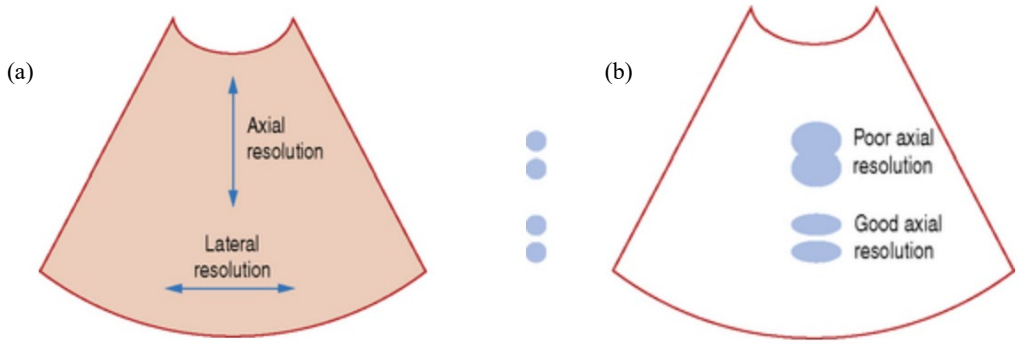


Fig. 1 (a) Difference between axial and lateral resolution [4] (b) Difference between poor and good axial resolution [4]

3.4 Transducers

The transducer is made out of a piezoelectric material which operates by changing the piezoelectric crystal's shape in response to an electric field, therefore exerting mechanical waves. This phenomenon is called piezoelectric effect [3]. A voltage is applied at a particular current and frequency to deform that crystal and propagates the sound. After that occurs, the sound waves will come back and hit the surface of that crystal which deforms it and creates a voltage across it that is later measured. The waves originally sent are decomposed into harmonic frequencies upon hitting tissues. These harmonics are later listened to instead of that original centre frequency, as they reduce noise and clutter, giving better lateral resolution.

4. Robotics for Positioning US Scanner

Robotics in Ultrasound scanning mainly consist of a robotic system that connects to a computer and a US station that holds the probe positioned on the patient's body. This section shows different types of robots used in US as well as several forms of applications. It starts with a quick review of physical requirements for probe positioning to obtain high quality pictures.

4.1 Probe Positioning Review

The probe's positioning is crucial in order to obtain an optimal image quality. Two main factors should be taken into consideration, the contact force F_c between the probe and the patient, and the probe's orientation. As mentioned previously, a better quality is obtained when the angle between the sound waves and the normal direction of the surface (θ_{AOI}) is close to zero, which would ensure that most echoes are reflected back to the transducer.

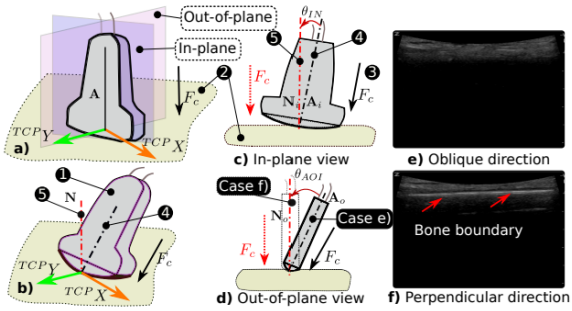


Fig. 2 [5] Impact of the probe orientation on US images. (a) and (b) describe ideal and non-ideal probe orientation in 3D respectively. (c) and (d) are in-plane and out-plane views respectively. (e) and (f) are images from ideal and non-ideal orientation

The central axis of the probe **A** and the surface normal **N** need to be aligned. These vectors can be decomposed into two orthogonal parts. First, an in-plane component $\mathbf{A_i}$ and $\mathbf{N_i}$ (Fig.2 c)) aligned with the front view of the probe which corresponds to its wider side. Secondly, an out-plane component $\mathbf{A_o}$ and $\mathbf{N_o}$ (Fig.2 d)) which is orthogonal to the in-plane component, or in other words aligned with the thinner sides

of the probe [5]. Fig.2 e) and f) show US images obtained when $\theta_{AOI} = 10^{o}$ and $\theta_{AOI} = 0^{o}$ respectively. The latter has a clear bone boundary which isn't depicted by the other scan. Thus the importance of having the angle of incidence close to zero. As for the contact force F_c, it is usually set to 3 – 15 N. If F_c is too large, the image quality declines due to deformations.

Z. Jiang et al. presented a full orientation optimization algorithm for better image quality. It focuses on aligning the probe to the surface's normal at the contact point both to the in- and out-plane. The US confidence map is exploited in order to optimize the in-plane orientation and identify its normal $\mathbf{N_i}$. It helps assess the image quality by estimating the signal loss. The matrix $C \in \mathbb{R}^2 \rightarrow [0,1]$ is divided into pixels holding the values of 1 meaning a perfect signal quality or 0 meaning no signal. Each value of C corresponding to each pixel represents the chance that the signal will reach the transducer. That method is therefore indicative of the quality of the image as the probe moves, showing what areas of the image are reflected back properly. After that, some calculations are made to finally come up with a fixed, optimal position for the probe. As for the out-plane orientation, $\mathbf{N_o}$ is computed using force values recorded during a fan motion [5]. However, several steps have to be taken to eliminate noise including a Kalman filter and a low pass filter. After that, the local minima registered whenever **A** and **N** are aligned are recorded. Only three of those (the ones that are most representative of the collected data) are later fed into a fusion algorithm to calculate the normal direction $\mathbf{N_o}$, which returns a corresponding $\mathbf{A_o}$ [5].

Consequently, the positioning of the probe is not to be taken lightly, as it highly affects the quality of the obtained images. Therefore, with robotics taking over and handling the probe, it is important to study the transducer's orientation.

4.2 Mountable Robots for US

Ultrasound is a user dependant tool and requires professional sonographers. From localizing the desired area of interest to holding the probe manually and applying the correct pressure, all while watching the US station screen and adjusting multiple settings accordingly, the sonographer has a huge responsibility in carrying the examination out. Robotic arms have been developed to assist radiologists in their examination. These robots efficiently handle the probe and precisely track the regions of interest (ROIs), and orient the transducer in an optimal way to obtain US images with a high level of accuracy. This can lead to the generation of uniformly spaced slices, which enables the generation of 3D images [6]. In fact, these require the integration of 2D images taken from different planes of the ROI with precisely known locations and orientations, which is achievable thanks to these robotic systems [6].

Unger et al. presents a way to autonomously track a vessel structure in US images in order to adjust the transducer's position. A KUKA LBR Med was used to dynamically position the US scanner. Fig. 3 shows the robotic arm holding the US probe and placed above an Agar-based phantom with an artificial anatomical structure to evaluate the tracking performance.

Fig. 3 [7] The KUKA LBR Med robotic arm

The KUKA LBR Med is a robotic arm designed to be used in several medical applications. For US scanning, the probe is connected to its flange and employed. It is flexible and allows for an easy automatic manoeuvre than can also be done manually if needed. Fig. 4 presents technical data on the following machine.

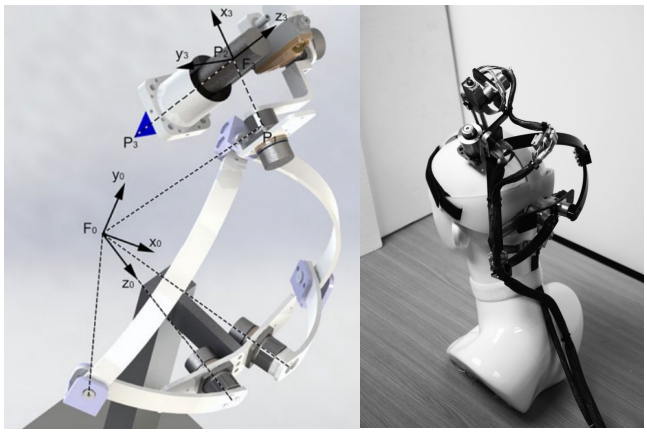
LBR Med	LBR Med 7 R800	LBR Med 14 R820
Max. total payload	7 kg	14 kg
Number of axes	7	7
Wrist variant	In-line wrist	In-line wrist
Mounting flange A7	DIN ISO 9409-1-A50	DIN ISO 9409-1-A50
Installation position	any	any
Positioning accuracy (ISO 9283)	±0.1 mm	±0.15 mm
Axis-specific speed accuracy (at max. speed)	±2 %	±2 %
Weight	25.5 kg	32.3 kg
Protection rating	IP54	IP54

Fig 4 [8] Technical data on the LBR Med 7 R800 and the LBR Med 14 R820

4.3 Wearable Robots for US

Although mountable robots are widely used for US applications, they present some difficulties when applied to certain body parts. Particularly, brain simulations require the patient to sit still during the examination for a long period of time in order to maximize positioning accuracy which can be easily distorted if the patient moves [9]. Additionally, non-invasive methods for brain simulations should be applied. There are several ones used nowadays, of which Transcranial Direct Current Simulation (tDCS) and Transcranial Magnetic Simulation (TMS) [9].

Junwoo et al. developed a wearable device for non-invasive transcranial US simulations. The device has a two DOF parallel spherical five-bar linkage that locate the simulator around the head. It also has a three DOF serial arm, exerting a five-DOF motion on a human head. The latter is used to control the simulation angle and depth. The robot used TMS which involves electric currents being induced at the ROI by applying a short, strong magnetic field by means of a magnetic coil. A low frequency ultrasound (<1MHz) is applied due to the presence of the skull which absorbs and refracts sound waves. The model is shown in Fig. 5 with its motion planes and positioned on a human head.



(a) (b)

Fig 5 [9] (a) Final design of the wearable robotic positioning system; (b) Final prototype

4.4 Future Directions Using Soft Robotics

The development of robots for US scanning purposes is limited for their rigid nature and hard mechanisms. This affects their flexibility and makes it complex for engineers to ensure compliance. To solve this problem, the creation of soft robots has been widely exploited. Soft robotics is a field that involves the design and fabrication of robots made out of compliant material. This allows the production of robots that can take a wider variety of shapes and perform more tasks which makes medical procedures easier and more efficient.

H. Ren et al. presented an attachable and portable soft robot that is able to optimize probe positioning and the contact force. The device (Fig. 6) consists of three main parts. The suction cups are used to attach the device to the body using a vacuum suction force of 50 N. The support structure is composed of a top plate that holds one soft pneumatic actuator in each of four quadrants, and a bottom plate which also has four quadrants that work independently and holes that attach it to the suction part. Finally, the soft pneumatic actuators which are used to steer the US probe.

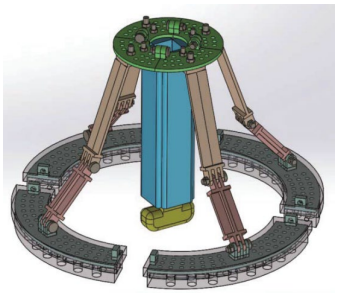


Fig. 6 [10] CAD model of the prototypical robotic system

Several active soft actuation technologies have been used to ensure flexibility, accuracy and safety in soft robotics, one of which is Flexible Fluid Actuators (FFAs) [11]. These structures are operated by a fluid and don't need a direct electric source. This allows them to be applied in the presence of radioactive materials and magnetic fields. They are able to move using a small number of parts, and offer safe interactions with tissues thanks to their compliance. Another technology for soft actuators are Shape Memory Alloys (SMAs) [11]. These smart materials have properties like high corrosion resistance, biocompatibility and non-magnetic behaviour. When heated, they bend and after a certain amount of time the material returns to its original shape. Therefore, it is said to be in its super elastic form. These devices are concentric tubes which are able to reach remote locations due to their bending properties. This allows for better reaching of small-diameter catheters. Shape Memory Polymers (SMPs) [12] are similarly used, undergoing shape change in response to heat. However, they exert lower actuation forces than SMAs and can't reset their shape on their own.

A particular type of soft devices are origami robots. Made out of flat composite sheets, they can be folded into any complex 3D morphology [12]. The tilings are separated by compliant joints that form a crease structure. Like any other robotic system, origami devices require actuation, sensing, computation and power. Because they contain many folds, their actuation sometimes requires a coordinated movement of subsets of these folds. SMAs can be used for direct actuation as well as SMPs such as polyolefin polystyrene and polyvinyl chloride which can be folded at the centimetre to millimetre scale [12]. Many robots are actuated with open-loop, ON/OFF control without sensor feedback. However, feedback control through embedded sensors is necessary to fabricate fold angles. For instance, strain sensors that measure bending using layers of microfluidic channels filled with metals that are liquid at room temperature. Moreover, sensors can be integrated to allow operation of the robot. For example, piezo resistive sensors can detect

loads in origami robot components. As for control, it entails low-level joint angle control which is realized in an open-loop with feedback control to regulate fold angles as mentioned previously. Additionally, the coordination of several joint actuators is required. Not to mention task-level planning that controls the robot's behaviour. A trade-off is to be made between how complex the folds are and the complexity of achievable shapes. Therefore, the shape of the original origami sheet should be optimized for the desired 3D shape. That way, it minimizes hassle and allows for easier control.

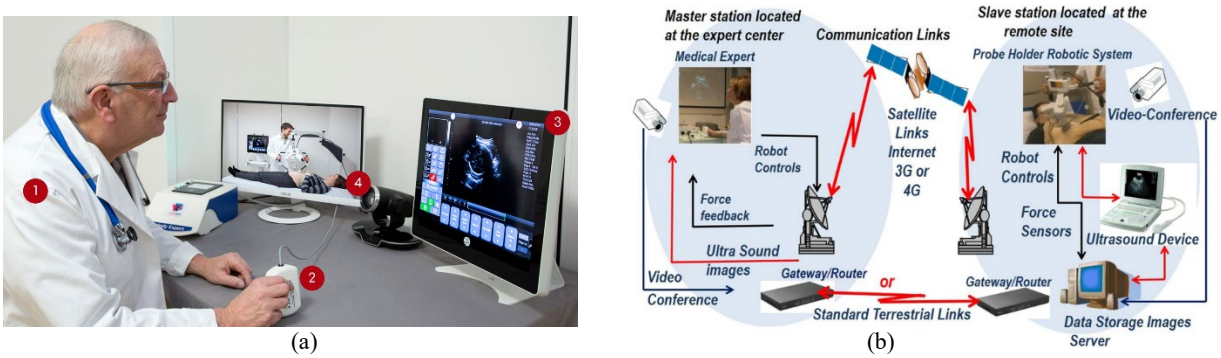
5. Applications of Robotics in US

After a detailed review on what types of robots are used in US scanning, this sections shows how US procedures are carried out. Although most require the assistance of a radiologist to monitor the examination and control the probe, methods for autonomous systems have recently been investigated.

5.1 Human Assisted Systems

Teleoperation has been put into practice for a while now. It simply means the operation of a machine from distance. In US, the patient is placed in a room with the robotic system that handles the probe which is initially positioned on his/her body by a nurse. The operation is then carried by a radiologist that controls the probe remotely and watches screen that shows the ultrasound images and another one that shows him the patient. That way, the radiologist doesn't have to be physically present to examine the patient. This method has several advantages, one of which is the increase in availability of doctors which is sometimes limited by distance [13]. Another relatable point is reducing the risk of virus transmission and contamination. In fact, tele-operated US scans have specifically been applied during the COVID-19 pandemic [15], to reduce the contact with patients. There are two mainly used robotic systems for teleoperations: MELODY and MGIUS-R3 [13].

MELODY is a robotic device fitted with an ultrasound probe connected to a US machine. It is mounted on a support arm. The nurse positions the device on the patient's body and follows the offsite radiologist's indications. The entire examination is conducted via video conference, allowing the radiologist to communicate with the patient and see the manipulations as they are being performed (Fig b). The robot reproduces exactly the same movement made remotely by the radiologist and follows the probe's angles of inclination in real time. The radiologist sees both the procedure and the US images on his screen (Fig. 7 a)).



(a) (b)

Fig. 7(a) Sonographer performing US scanning remotely [16]; (b) Robotized tele-ultrasound using the MELODY system [14]

MGIUS-R3 (Fig. 8) is another robotic system used specifically for the abdomen, small parts, vascular, urology, gynaecology and musculoskeletal [17]. Similar to MELODY, it has a real time transmission and offers high quality image. It includes a multi-sensor fusion robotic arm with a scanning diameter of 850mm

and 6 degrees of freedom. Just like MELODY, it reproduces the expert's movements. The robotic probe handled by the latter has a total of 6D data that can be collected: 3D rotation, 2D plane and a 1D force.

Fig. 9 [17] The MGIUS-R3 robotic system

5.2 Autonomous Systems

In hopes to reduce human intervention, several methods have been explored in autonomous US systems. Each have distinct applications and facilitate the sonographer's mission.

von Haxthausen et al. discussed autonomous image acquisition highlighting three specific notions: 3D image reconstruction, autonomous trajectory planning and probe positioning, and optimizing image quality. In order to acquire volumetric images, [18] shows a depth camera that captures relative positions of the patient and autonomously plan a scan path for the US robot. Pre-recorded images such as CT or MRI were later used to evaluate the trajectory needed to image the desired volume of interest. To improve image quality, autonomous probe positioning systems are developed so that the in- and out-planes are automatically adjusted in an optimal way (Fig. 2).

Goel R. et al. designed an autonomous US scanning system that finds the optimal area to scan and performs the actual scanning. Bayesian optimization is used to guide the probe to several points on the unknown surface. Regions with high vessel density are depicted in order to narrow the area of interest. But since veins and arteries randomly move along the body, a limited number of samples is collected. A Gaussian Process Regression (GPR) is therefore implemented to estimate the unknown function based off of the collected data. It works by computing a mean function and standard deviation function to predict the value and its confidence at any given point respectively. Then, Bayesian optimization relies on GPR to find the function estimate's distribution and uses an acquisition function to compute the best next location to sample from. After detecting the area of interest, a force based controller is employed to scan the surface. Fig.10 illustrates the entire procedure from identifying the targeted region to scanning it.

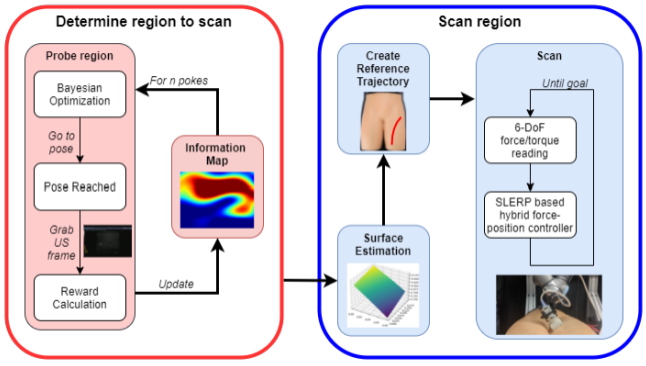


Fig. 10 [19] High-level pipeline depicting the proposed autonomous scanning framework

6. Localization of Targeted Areas in US Scanning

Localizing targeted areas of interests has always been up for research. Artificial Intelligence (AI) is a key concept for determining the exact ROI to scan. The algorithm takes in live images of the desired area as inputs using a camera attached to the US probe for example. After a series of operations, it compares the images to predisposed pictures of the same tissue used for the same scanning purposes and predicts the exact scanning area. In this sections, an overview of a basic convolutional neural network (CNN) is given in order to understand how AI algorithms actually work, followed by a review of previous studies that involve the use of AI for localization purposes.

6.1 Overview of Basic CNN Concept

CNN is a network architecture for deep learning that learns directly from data and is able to find patterns in image in order to recognize objects. It is comprised by convolution layers which are repeated a certain amount of time until the classification part comes to play. A convolution layer consists of convolution operation followed by Rectified Linear Unit (ReLU) and a pooling stage.

First, an image, or a three dimensional matrix is fed to the algorithm. Fig. 11 shows how an input image looks like. It has the following dimensions (height, width, colour channels). The numbers in each pixel (entry) represent their corresponding colour. For instance, in a black and white images, each pixel can only hold values around 0 (for black) and 1 or 255 (for white).

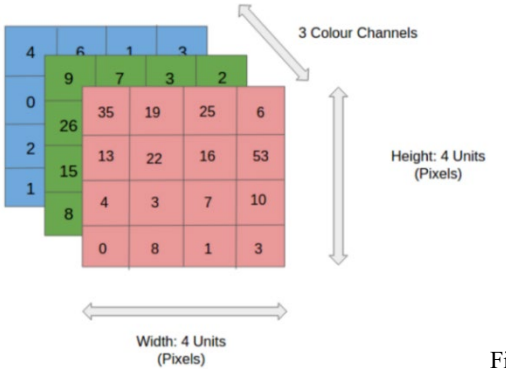


Fig. 11 [21] A 4 x 4 x 4 RGB (Red-Green-Blue) image

The input matrix is then convoluted with a kernel (or filter). This is done by overlapping the kernel with all the different parts of the input matrix, multiplying each overlapping entries together and calculating their overall sum.

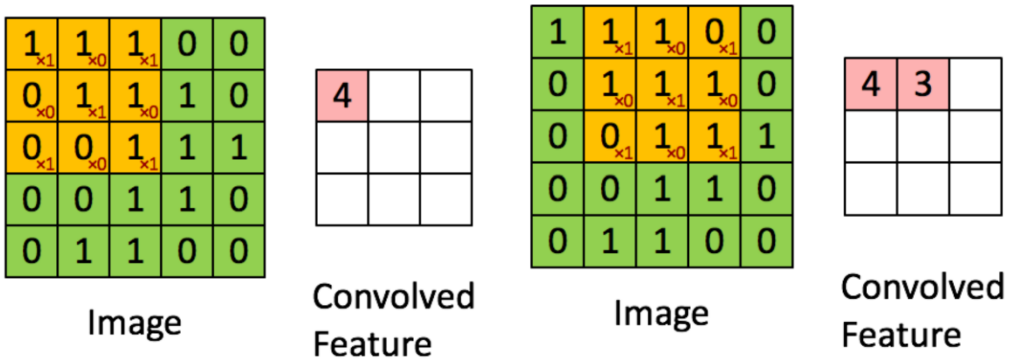


Fig. 12 [21] Convolution of a 5 x 5 x 1 with a 3 x 3 kernel and resulting convolved feature

Fig. 12 shows a (5 x 5 x 1) input image convoluted with the following (3 x 3) kernel $\begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}$.

The value of the first entry is calculated as such: 1x1 + 0x0 + 0x1 + 0x1 + 1x1 + 0x0 + 1x1 + 1x0 + 1x1=4. The other entries are similarly computed until we obtain a final output:

$$\begin{pmatrix} 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 \end{pmatrix} * \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 4 & 3 & 4 \\ 2 & 4 & 3 \\ 2 & 3 & 4 \end{pmatrix}$$

For 3D images, the convolution is evaluated in the same manner but each entry in the final output is the sum of all three numbers resulting from every convolution of each one of the 3 colour channels' matrices with the kernel.

The stride represents the portion of the image over which the kernel is hovering (i.e number of shifts) [21]. In the above example, the stride length was equal to 1 since the kernel is not jumping over any column. If the stride was equal to 2, we would have ended up with 1 (2 x 2 x 1) matrix.

This operation is done several times in order to end up with a simplified image that still holds all of its original form's key features which represent the pictured object. The first ConvLayer is responsible for capturing the low-level features such as edges, colour, gradient orientation, etc. With additional layers, it gradually depicts the high-level features as well, giving us a networks which has a wholesome understanding of the image's dataset [21].

Moving along the convolutional layer, a ReLU is used to eliminate any negative number found in the convolved matrix. The activation function applied is $f(x) = \max(0, x)$, therefore changing any negative number to zero. Other functions can be used like the sigmoid function $\sigma(x) = (1 + e^{-x})^{-1}$ and the saturating hyperbolic tangent $f(x) = \tanh(x)$ or $f(x) = |\tanh(x)|$ (Fig. 13).

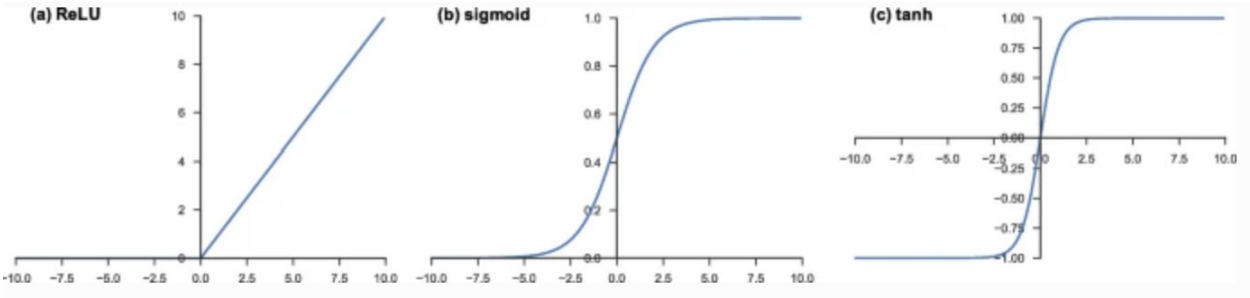


Fig. 13 [20] Activation functions commonly applied to neural networks: a) rectified linear unit; b) sigmoid; d) hyperbolic tangent

After that, a pooling layer takes place. It is used to reduce the special size of the convolved feature by eliminating the values that don't have a significant impact on the image and keeping the predominant ones. There are two types of pooling: Max pooling, which only takes the largest value in a block, and an Average pooling which evaluates the mean of all values (Fig. 14).

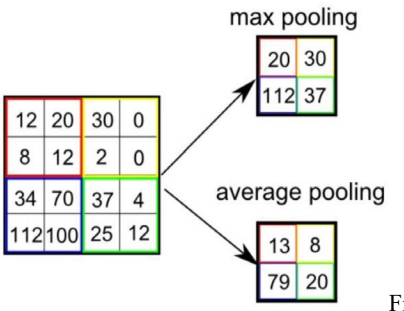


Fig. 14 [21] Types of pooling: max pooling and average pooling

The following image clearly illustrates how US pictures change after max pooling.

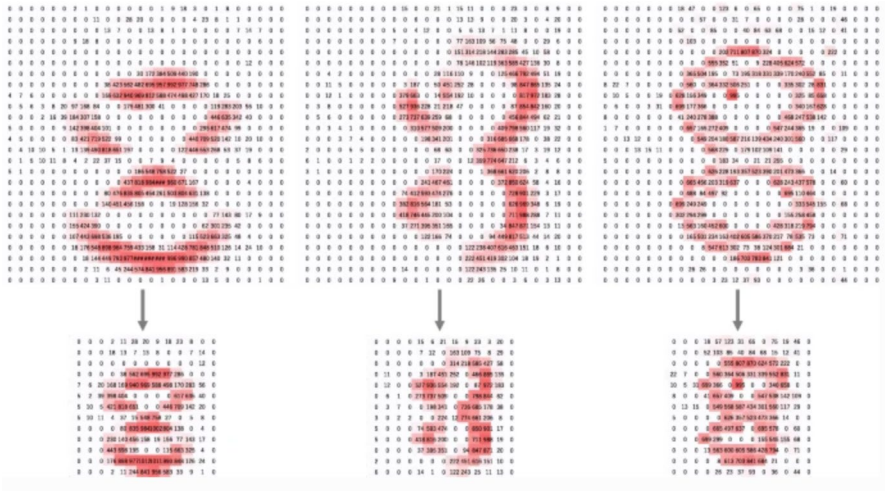


Fig. 15 [20] Examples of the max pooling operation. Note that images in the upper row are down sampled by a factor of 2, from 26 × 26 to 13 × 13

Finally, the output feature map of the final convolution or pooling layer is flattened into a 1-dimensional array and fed to a Fully Connected Layer (FCL) which includes pre-saved pictures of the same object. The FCL will map all these pictures and output a vector of K dimensions where K is the number of classes the network will be able to predict [20]. The vector contains the probabilities of each class of any classified image. After that, a final layer activation function (like softmax) is applied to provide the classification output. A summary of the CNN model is illustrated in Fig. 16.

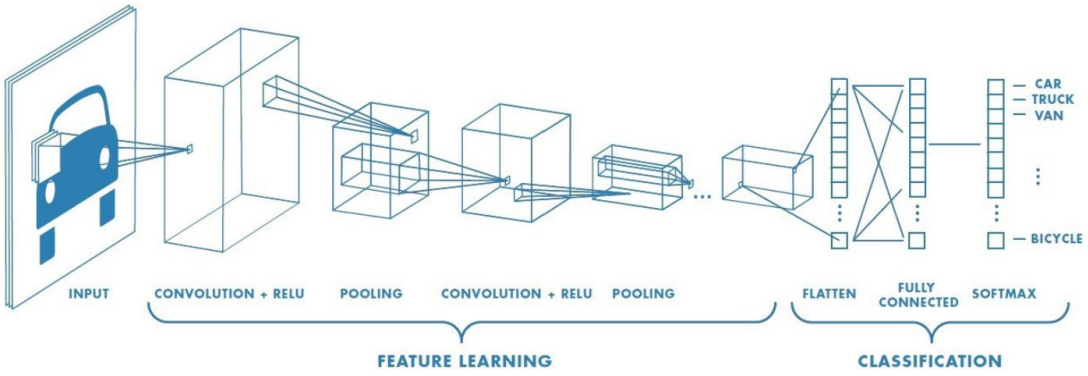


Fig. 16 [21] CNN model from convolutional layers to classification in fully connected layers

6.2 AI Algorithms Used for US Applications

Zhu J et al. constructed a deep CNN model which helps in localizing and classifying thyroid nodules in order to help treat thyroid cancer. The model is called Brief Efficient Thyroid Network (BETNET). It is composed of 16 convolution layers, 3 fully connected layers and the softmax layer. Fig. 17 shows the layout of the BETNET with information about specific operations and output sizes. The model was trained with 1000 thyroid nodule pictures and then by increasing the number of training set images. It was successful in accurately identifying thyroid cancer and localizing nodules. Fig. 18 shows the obtained US images with the nodules framed in an orange box. By increasing the number of images in the training set, accuracy improved considerably (from 76.5% to 97.8%)

Layer name	Output size	Activation	Layer design
conv_block1_1	112×112×64	ReLU	3×3 conv, 64
conv_block1_2	112×112×64	ReLU	3×3 conv, 64
Pool_1	56×56×64		
conv_block2_1	56×56×128	ReLU	3×3 conv, 128
conv_block2_2	56×56×128	ReLU	3×3 conv, 128
Pool_2	28×28×128		
conv_block3_1	28×28×256	ReLU	3×3 conv, 256
conv_block3_2	28×28×256	ReLU	3×3 conv, 256
conv_block3_3	28×28×256	ReLU	3×3 conv, 256
conv_block3_4	28×28×256	ReLU	3×3 conv, 256
Pool_3	14×14×256		
conv_block4_1	14×14×512	ReLU	3×3 conv, 512
conv_block4_2	14×14×512	ReLU	3×3 conv, 512
conv_block4_3	14×14×512	ReLU	3×3 conv, 512
conv_block4_4	14×14×512	ReLU	3×3 conv, 512
Pool_4	7×7×512		
conv_block5_1	7×7×512	ReLU	3×3 conv, 512
conv_block5_2	7×7×512	ReLU	3×3 conv, 512
conv_block5_3	7×7×512	ReLU	3×3 conv, 512
conv_block5_4	7×7×512	ReLU	3×3 conv, 512
Reshape	25,088		
FC_1	4,096	ReLU	250,888×4,096
FC_2	4,096	ReLU	4,096×4,096
FC_3	2	ReLU	4,096×2
Softmax	2		

conv_block, convolution block; ReLU, rectified linear Unit; FC, fully connected.

Fig. 17 [22] Structure of the BETNET model

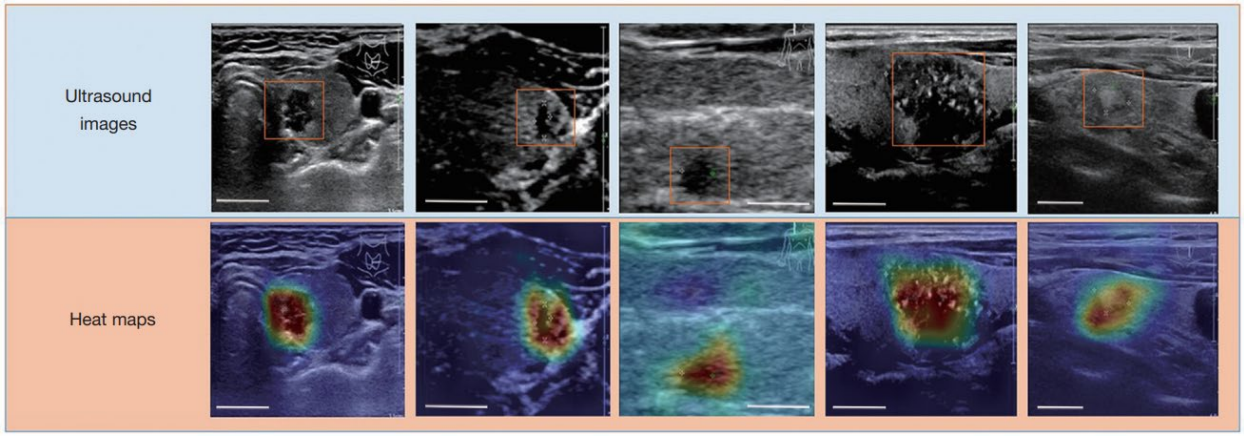


Fig.18 [22] Localization of thyroid nodules. The orange box area is the location of the thyroid nodules, and the heat map represents the visualization results. Each column shows the same ultrasound image. The warm tone region of the visualization image is the most important part that the neural network could recognize. By contrast, the cold tone region is the least important feature. The white scale bars represent 1 cm on the ultrasound images.

M, Pesteie et al. developed another CNN model to localize the needle target for epidural needle injection. The epidural space refers to the laminae of vertebrae's base positioned between ligamentum flavum and dura mater in the human spine. The network architecture consists of three convolutional layers. The first one has ten 5 x 5 kernels and the other two have ten 3 x 3 kernels, each convolution is followed by a max pooling layer of 2 x 2 pool size. The last convolutional layer has feature maps with multi-scale which are augmented to provide a better representation of the spinal anatomy. Finally, a fully connected layer with 2000 neurons (connections) follows the last pooling layer. Therefore, the augmented feature maps are classified. The training data set comprised 501,298 image patches which included as much target as non-target images. The model succeeded in localizing the anterior base of the lamina which is what the needle should be targeting.

7. Conclusion

In conclusion, ultrasound has opened doors for numerous applications in the medical field. It is used to monitor organs and tissues, detect tumours or infections, or even target specific locations for drug delivery. None of these practices would have been made possible without the fabrication of robotic systems that help guide these procedures. Robotic arms like MELODY or MGIUS-R3 with mounted probes are just a few examples of how electronic devices can help with that. With the development of wearable robots and origami shaped devices, designs are limitless, and number of medical applications are expending. Comfort in US scanning operations is also increasing for both doctors and patients. They can be carried out remotely thanks to teleoperation, eliminating the need for the sonographer to be present on site. Human assistance is slowly fading away for certain tasks with US scanning moving towards autonomous functioning. On a slightly different note, AI is at its peak when it comes to localizing areas to be scanned. CNNs are used to predict the exact desired targeted area by comparing it to a set of images that were initially fed to the algorithm, and therefore predicting the correct ROI. This improves accuracy immensely and allows for better and safer US applications.

8. References

[1] Long, F. Lim, I. Sum, T. “Collapsible Mount for Transcranial Doppler (COMTRAD). BN5101 Final Report. National University of Singapore.

[2] Sigrist, Rosa & Liau, Joy & El Kaffas, Ahmed & Chammas, Maria & Willmann, Jürgen. (2017). Ultrasound Elastography: Review of Techniques and Clinical Applications. Theranostics. 7. 1303-1329. 10.7150/thno.18650.

[3] Coatney, R. (n.d.). *Ultrasound imaging: Principles and applications in rodent research*. ILAR journal. Retrieved August 13, 2022, from https://pubmed.ncbi.nlm.nih.gov/11406722/

[4] *Resolution*. Radiology Key. (2016, March 10). Retrieved August 13, 2022, from https://radiologykey.com/resolution/

[5] Z. Jiang et al., "Automatic Normal Positioning of Robotic Ultrasound Probe Based Only on Confidence Map Optimization and Force Measurement," in IEEE Robotics and Automation Letters, vol. 5, no. 2, pp. 1342-1349, April 2020, doi: 10.1109/LRA.2020.2967682.

[6] Teye, J. (1970, January 1). *3D robotic ultrasound system: A 3D volumetric ultrasound imaging system using a CMUT phased array and a robotic arm*. TU Delft Repositories. Retrieved August 13, 2022, from http://resolver.tudelft.nl/uuid:ced96a5f-8174-473b-a20d-298804db712b

[7] Unger, Michael, Berger, Johann, Gerold, Bjoern and Melzer, Andreas. "Robot-assisted Ultrasound-guided Tracking of Anatomical Structures for the Application of Focused Ultrasound" *Current Directions in Biomedical Engineering*, vol. 6, no. 3, 2020, pp. 123-126. https://doi.org/10.1515/cdbme-2020-3032

[8] *LBR MED: A collaborative robot for Medical Technology*. KUKA AG. (n.d.). Retrieved August 13, 2022, from https://www.kuka.com/en-de/industries/health-care/kuka-medical-robotics/lbr-med

[9] Kim, Joonwoo & Lee, Sungon. (2016). Development of a Wearable Robotic Positioning System for Noninvasive Transcranial Focused Ultrasound Stimulation. IEEE/ASME Transactions on Mechatronics. 21. 1-1. 10.1109/TMECH.2016.2580500.

[10] H. Ren, X. Gu and K. L. Tan, "Human-compliant body-attached soft robots towards automatic cooperative ultrasound imaging," 2016 IEEE 20th International Conference on Computer Supported Cooperative Work in Design (CSCWD), 2016, pp. 653-658, doi: 10.1109/CSCWD.2016.7566066.

[11] Cianchetti, Matteo & Laschi, Cecilia & Menciassi, Arianna & Dario, Paolo. (2018). Biomedical applications of soft robotics. Nature Reviews Materials. 3. 10.1038/s41578-018-0022-y.

[12] Rus, Daniela & Tolley, Michael. (2018). Design, fabrication and control of origami robots. Nature Reviews Materials. 3. 1. 10.1038/s41578-018-0009-8.

[13] von Haxthausen, Felix & Böttger, Sven & Wulff, Daniel & Hagenah, J. & Garcia-Vázquez, Veronica & Ipsen, Svenja. (2021). Medical Robotics for Ultrasound Imaging: Current Systems and Future Trends. Current Robotics Reports. 2. 10.1007/s43154-020-00037-y.

[14] Avgousti, Sotiris & Christoforou, Eftychios & Panayides, Andreas & Voskarides, Sotos & Novales, Cyril & Nouaille, Laurence & Pattichis, C. & Vieyres, Pierre. (2016). Medical telerobotic systems: Current status and future trends. BioMedical Engineering OnLine. 15. 10.1186/s12938-016-0217-7.

[15] Akbari Mojtaba, Carriere Jay, Meyer Tyler, Sloboda Ron, Husain Siraj, Usmani Nawaid, Tavakoli Mahdi. "Robotic Ultrasound Scanning with Real-Time Image-Based Force Adjustment: Quick Response for Enabling Physical Distancing During the COVID-19 Pandemic". Frontiers in Robotics and AI. Vol. 8. 2021. https://www.frontiersin.org/articles/10.3389/frobt.2021.645424

[16] *Melody, a remote, robotic ultrasound solution*. AdEchoTech. (n.d.). Retrieved August 13, 2022, from https://www.adechotech.com/products/

[17] *MGIUS-R3 robotic ultrasound system-MGI-leading life science innovation*. MGI. (n.d.). Retrieved August 13, 2022, from https://en.mgi-tech.com/products/instruments_info/11/

[18] Q. Huang, J. Lan and X. Li, "Robotic Arm Based Automatic Ultrasound Scanning for Three-Dimensional Imaging," in IEEE Transactions on Industrial Informatics, vol. 15, no. 2, pp. 1173-1182, Feb. 2019, doi: 10.1109/TII.2018.2871864.

[19] Goel, R., Abhimanyu, F., Patel, K., Galeotti, J., & Choset, H. (2022, May). Autonomous Ultrasound Scanning using Bayesian Optimization and Hybrid Force Control. In *2022 International Conference on Robotics and Automation (ICRA)* (pp. 8396-8402). IEEE.

[20] Yamashita, R., Nishio, M., Do, R. K. G., & Togashi, K. (2018, June 22). *Convolutional Neural Networks: An overview and application in radiology - insights into imaging*. SpringerOpen. Retrieved August 13, 2022, from https://insightsimaging.springeropen.com/articles/10.1007/s13244-018-0639-9#Sec2

[21] Saha, S. (2018, December 17). *A comprehensive guide to Convolutional Neural Networks-the eli5 way*. Medium. Retrieved August 13, 2022, from https://towardsdatascience.com/a-comprehensive-guide-to-convolutional-neural-networks-the-eli5-way-3bd2b1164a53

[22] Zhu J, Zhang S, Yu R, Liu Z, Gao H, Yue B, Liu X, Zheng X, Gao M, Wei X. An efficient deep convolutional neural network model for visual localization and automatic diagnosis of thyroid nodules on ultrasound images. Quant Imaging Med Surg. 2021 Apr;11(4):1368-1380. doi: 10.21037/qims-20-538. PMID: 33816175; PMCID: PMC7930675.

[23] M. Pesteie, V. Lessoway, P. Abolmaesumi and R. N. Rohling, "Automatic Localization of the Needle Target for Ultrasound-Guided Epidural Injections," in IEEE Transactions on Medical Imaging, vol. 37, no. 1, pp. 81-92, Jan. 2018, doi: 10.1109/TMI.2017.2739110.