A Semi-Autonomous Robotic System for Remote Trauma Assessment

B. Mathur[1*], A. Topiwala[1*], S. Schaffer[1], M. Kam[1], H. Saeidi[1], T. Fleiter[2],
and A. Krieger[1], *Member, IEEE*

Abstract— Trauma is among the leading causes of death in the United States with up to 29% of pre-hospital trauma deaths attributed to uncontrolled hemorrhages. This paper reports a semi-autonomous robotic system capable of assessing trauma using 2D and 3D image analysis and enabling remote focused assessment with sonography for trauma (FAST) *en route* to the hospital for earlier trauma diagnosis and faster initialization of life saving care. The system was able to accurately calculate FAST scan positions of patient specific phantoms using the measured phantom sizes and positions of the umbilicus. The system was capable of accurately classifying and localizing wounds, so they can be avoided during the ultrasound scan. These objects were localized with an accuracy of 0.94 ± 0.179cm and FAST exam locations were estimated with an accuracy of 2.2 ± 1.88cm. A radiologist successfully completed a remote FAST scan of the phantom using the system with improved image quality over manual scans, demonstrating feasibility of the system.

I. INTRODUCTION

Trauma was ranked the third leading cause of death in all age groups and ranked first for people 1-44 years old. Injuries related to falls, firearms, and motor vehicle accidents are the most common causes of mortality within this group [1]. The Golden Hour concept established that the chances of survival for a severe trauma victim diminish rapidly with passing time until the patient's injuries, specifically hemorrhage, are addressed and stabilization is achieved [2]. Despite the development of coordinated statewide Emergency Medical Systems (EMS) providing optimized initial care in the field and rapid transport to a dedicated Trauma Center [3], pre-hospital mortality remains a significant problem. Up to 29% of the pre-hospital trauma deaths could be classified as survivable and were mainly attributed to uncontrolled abdominal hemorrhages [4].

The Focused Assessment with Sonography for Trauma (FAST) exam was developed for diagnosing abdominal hemorrhages [5]. FAST is a relatively simple ultrasound imaging method to identify significant hemorrhage in the peritoneal, pericardial, and pleural spaces as well as to identify lung collapse (pneumothorax) [6]. The technique focuses on four major scan locations as shown in Figure 2(e). Despite the proven efficacy of the FAST scan and the availability of compact and mobile ultrasonography systems, the use of

*B. Mathur and A. Topiwala contributed equally to this work.

[1] B. Mathur, A. Topiwala, S. Schaffer, M. Kam, H. Saeidi, and A. Krieger are with the Mechanical Engineering Department, University of Maryland, College Park, MD 20742, USA. E-mail: {bmathur, anirudht, saul, mkam, hsaeidi, axel}@umd.edu.

[2] T. Fleiter is with the University of Maryland, School of Medicine, Department of Diagnostic Radiology, R.Cowley Shock Trauma Center, Baltimore, MD 21201, USA. E-mail: tfleiter@umm.edu
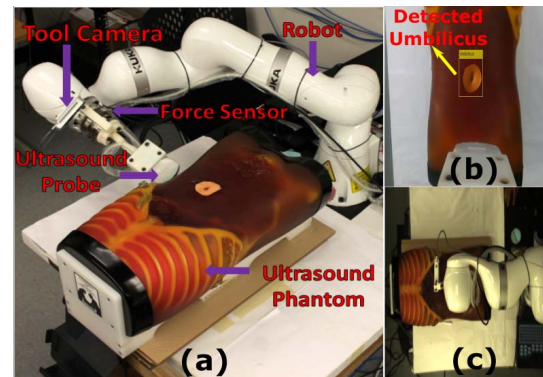
Fig. 1: a) Robotic system for remote trauma assessment. Examples of b) tool camera view with detected umbilicus landmark and c) world camera view.

FAST scans as a triage method has been hindered by the lack of emergency medical personnel trained in performing and interpreting the exam [7]. Furthermore, performing a FAST exam requires the undivided attention of the first responder and therefore interferes with tasks necessary to stabilize the patient, including mechanical hemorrhage control and cardiopulmonary resuscitation.

This paper reports the design of a novel, semi-autonomous robotic ultrasound system for remote trauma assessment and its evaluation on a phantom. The system has the following innovative capabilities: 1) autonomous 2D/3D torso scanning, 2) autonomous classification and 3D localization of torso landmarks (umbilicus) and wounds, 3) autonomous localization and robotic positioning of the ultrasound probe at FAST scan positions, and 4) remote, tele-manipulated robotic FAST scanning which uses Artificial Potential Fields (APF) to assist and guide a physician to avoid wounds. The system, once installed in an ambulance or emergency helicopter, enables a remote physician to conduct a FAST exam on a trauma patient *en route* to the hospital. The availability of the semi-autonomous FAST robotic system would not only transport essential diagnostic procedures from the hospital into the pre-hospital setting but also enable the development of novel robot-assisted methods for invasive hemorrhage control. The virtual presence of a physician in the ambulance with the trauma patient could have a significant impact on the pre-hospital mortality of trauma patients, especially if paired with aggressive hemorrhage control methods.

A. Prior Work

1) Robotic Ultrasonography: The wearable robotic system in [8] provides satisfactory images for diagnosis of
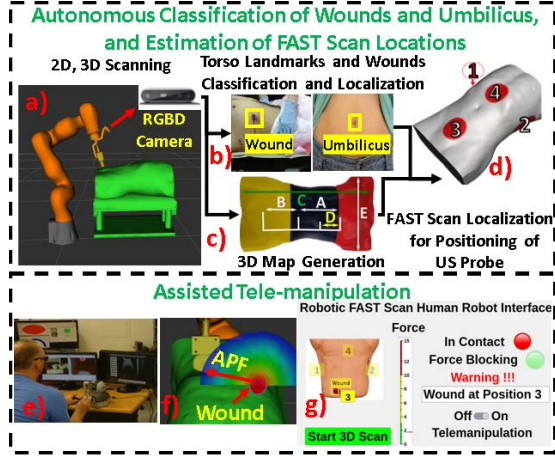
Fig. 2: Work-flow for semi-autonomous robotic diagnosis of trauma patients

hemorrhages. The system requires a bystander to attach and position the corset on the victim after which the robot is remotely-controlled by a physician to perform the FAST exam. The downside of this robot is that its use is limited by the extent of injuries of the victim. It also relies on the fact that these lifesaving units need to be installed at 300m intervals across the town to be easily accessible in case of an emergency. The cooperative controlled robotic system in [9] provides guidance to the users in the form of haptic feedback, utilizing a virtual fixture whose parameters are tuned by real-time US image feedback during the placement of the US probe. As this system is primarily designed to assist in radiotherapy, for guidance, it makes use of a reference US image of the anatomy of interest which is not available in case of a trauma victim. Even though such a cooperative robotic system with a virtual fixture makes it easier to perform FAST exam in an ambulance, it will require the undivided attention of a first responder. The autonomous system in [10] shows promising results in estimating the position of the epigastric region using visual landmarks such as the umbilicus and mammary papillae for autonomous scanning of the liver. However, additional features such as the ability to avoid wounds, and US based visual servoing to accurately find the anatomy of interest are required for such a system to be used for detection of internal injuries. The authors in [11] provide a comprehensive overview of existing robotic arm-assisted ultrasonography systems along with a list of technical requirements from such a system, and potential clinical applications where robotic systems could be advantageous in the future.

To our knowledge, our robotic system is the first semi-autonomous robotic ultrasound (US) system to incorporate all the necessary operational and safety features required in a system for remote assessment of trauma. These include autonomous estimation and positioning of the probe at US scan locations, ability to detect and possibly avoid unscannable regions (wounds), virtual fixture to prevent application of excessive forces on the patient, and assisted tele-manipulated FAST exam with haptic feedback performed by a remote physician.

2) Object Classification and Localization: In 2012, [12] won the ImageNet Large Scale Visual Recognition Challenge (ILSVRC) [13] using convolution neural networks (CNNs). Since then CNNs have become the gold standard to address the image classification and detection problem. In [14], CNNs were used to get substantial classification accuracy for melanoma on clinical images. This idea was extended to wound classification by [15]. They used deep neural networks (DNN) along with color thresholding to segment wounds from clinical images to analyze healing of wounds. The limitation of this work is that the clinical images used are zoomed into the wound, which is not the case in our application when scanning the entire torso. To address this, we need a detection algorithm which can detect wounds in the presence of high amounts of background information. The sliding window approach in [16] to solve this detection problem is both time consuming and memory intensive. The region-based convolution neural network (CNN) like Faster R-CNN [17] addressed this with its increased accuracy of detection and minimal memory usage. Although these algorithms have provided good classification and detection results with established multi-class generic object detection data sets like PASCAL VOC 2007 [18] and Microsoft COCO [19], they have not been optimized and tested for detecting umbilici or wounds on torso images, largely because of lack of an established datasets. To this extent, we propose an algorithm that will be able to carry out robust wound and umbilicus detection and localization, for images with significant background information using a custom large data set for both the classes.

II. METHODS

Fig. 2 shows the step-by-step work-flow of the remote trauma diagnosis process with our robotic system. The assessment begins with the robot autonomously sweeping over the patient to record several point cloud images, which are stitched together to construct a 3D image of the torso (Section II-B). 2D RGB images are also collected during this process for the detection and localization of the umbilicus and the wounds (Section II-C). The 3D image along with the position of the umbilicus are used to autonomously estimate the FAST scan locations (Section II-D).

The robot autonomously positions the ultrasound probe at a point close to each FAST scan position, before a remote physician assumes control to perform the assisted tele-manipulated ultrasound scan with haptic feedback (Section II-E). A Graphical User Interface (GUI) assists the remote operator in switching between FAST scan positions and help with the tele-manipulation.

In this section, we describe the components of the developed robotic system including the robot, imaging system, and control techniques used for the semi-autonomous remote assessment of trauma.

A. Testbed

Fig.1 shows a picture of our robotic system for remote trauma assessment, consisting of a) a 7 degrees of freedom

(DOF) medical light weight robot (LWR) (KUKA AG, Augsburg Germany), b) a world frame color camera Chameleon CMLN-13S2C (Sony, Tokyo, Japan) mounted above the patient and robot, c) an ultrasound probe, d) a RGB-D tool camera RealSense D415 (Intel, Santa Clara, USA) and e) a SI-65-5 six-axis F/T Gamma transducer (ATI Industrial Automation, Apex, North Carolina). The following describes the sub-components of our testbed.

1) Ultrasound system integration into the robotic system: The CMS600P2 Portable Ultrasound Scanner from Contec (Contec Medical Systems Co. Ltd, Hebei, China) was used with a 3.5 MHz convex probe for US imaging. A custom 3D printed mount was used to attach the US probe and tool camera to the force sensor. To maximize the ease of use of the system, the ultrasound probe was set to be coaxial with the last joint of the robot, the flange rotation. The tool camera was mounted behind the probe, providing both a view of the tool and the scene in front of the tool.

2) Phantoms: Two types of custom phantoms were developed to evaluate each sub-system's performance and a commercial phantom for FAST training was used for performing a complete trauma assessment with the fully-integrated robotic system.

a) Wound Phantom: A foam phantom (see Fig. 2(g)) that more closely resembles human skin tone as compared to the other two phantoms was used to validate the umbilicus and wound detection and localization subsystems. Artificial wounds were simulated using synthetic blood while pictures were used for umbilici. These were placed at various positions to carry out the detection and localization experiment as described in Section III-A.

b) FAST scan localization phantoms: To test the system's ability to localize and autonomously navigate to the four FAST scan locations, five patient specific localization phantoms were created. These phantoms are 3D printed versions of anonymized CT scans of real trauma patients, selected to provide a variety in size and sex. A radiologist extracted torso models from the CT scans and exported them into a mesh file format along with highlighting each patient's FAST scan positions. As similar FAST scans can be obtained by positioning the probe at a variety of skin points, circular regions (1cm diameter) instead of points were highlighted (see Fig 2(g)). The meshes were then denoised, smoothed, hollowed, and further segmented into three parts to fit on the 3D printer (Makerbot 5th Generation Replicator, Brooklyn, NY). Before printing, the highlighted FAST scan regions were recessed by 1mm so that the accuracy of the robotic system could be ascertained during testing. To speed up the printing process, the prints were paralleled across identical printers, leading to multi-colored phantoms. As the 3D point cloud and reconstruction are only used to estimate the size of the phantoms, their colors do not affect any of the results described in Sections III-B and II-D.

c) FAST Ultrasound Phantom: The FAST/Acute Abdomen Phantom (Kyoto Kagaku Co. Ltd, Kyoto, Japan) was used to test the fully-integrated system. This phantom features realistic internal hemorrhages at all four FAST locations

including at the pericardium and bilateral chambers, intra-abdominal hemorrhages around the liver, the spleen, and the urinary bladder.

3) Haptic feedback and control input device - Master: Geomagic Touch (3D Systems, South Carolina, USA), a widely used Haptic device, was used to provide bilateral tele-manipulation commands from the operator to the robot, receive haptic feedback proportional to the forces acting on the ultrasound probe, and apply APFs to guide the user as described in Section II-E.

4) Robotic software architecture: The control system including the planning algorithms, robot controllers, computer vision, and the control allocation strategies are integrated via Robot Operating System (ROS) [20]. Kinematics and Dynamics Library (KDL) in Open Robot Control Systems (OROCOS) [21] is used to transform the task-space trajectories of the robot to the joint space trajectories, which is the final output of the high-level autonomous control system. Finally, IIWA stack developed in [22] helps to apply the low-level controllers to the robot to follow the desired joint-space trajectories. The GUI (Fig. 2(g)) was developed in MATLAB and linked to ROS.

B. Autonomous 2D, 3D Scanning and 3D Map Generation

The robot performs a semicircular motion around the phantom at a distance of 30cm between camera and skin to collect point clouds and 2D images with the RGB-D camera in 21 pre-programmed positions. As the US probe blocks the lower half of the camera view, point clouds from each point are cropped to remove the obstructed region. With these multiple scenes, the system performs point cloud registration and reconstructs a 3D model by using the iterative closest point (ICP) algorithm [23], which is implemented in point cloud library (PCL) [24] and pointmatcher [25] following prior noise removal with a PCL passthrough filter [26]. We estimate the plane of the table using forward kinematics and crop the reconstructed phantom below this plane to remove the background table and extract the reconstructed torso.

C. Umbilicus and Wound Classification and Localization

2D RGB images are used to detect the umbilicus and wounds, which are then localized by transforming the pixels into the frame of the 3D reconstruction. The umbilicus is used as a visual anatomical landmark for estimating the FAST scan points on the phantom as discussed in Section II-D. Haptic feedback is generated using APFs emerging from the center of the wound in the direction of the EE to help guide the probe away from the wounds as described in Section II-E. The system also displays a warning on the GUI to notify the radiologist of any wounded regions near a FAST scan location (see Fig. 2(g)). Identifying severe wounds also helps the physician prioritize the order of locations in which the trauma assessment is conducted.

Faster R-CNN [17] for object detection was chosen as it simultaneously trains the region proposal, and the classification networks. This helps improve the region proposal quality and estimation as opposed to using fixed region proposals,
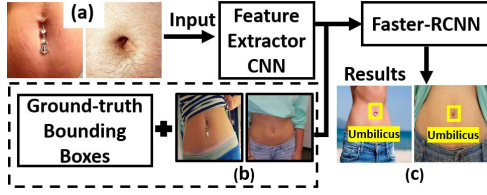
Fig. 3: Training and Detection Flowchart (a) Feature Extractor CNN Training Images. (b) Faster-RCNN Training Images. (c) Detection results on Detection Test Set.

such as in selective search [27]. It also has a faster detection time as compared to other region based CNNs.

1) Database Preparation: Data was generated for two classes: umbilicus and wounds. While some of the wound images were taken from the Medtec Wound Database (http://www.medetec.co.uk/files/medetec-image-databases.html), the rest were obtained from Google Images. The images were annotated and cross verified with an expert radiologist. To limit the amount of variation in the input data, wounds were restricted to gunshot wounds. To train the detector or the Faster-RCNN network, a trained feature extractor CNN or a classifier is required. Therefore, to train and validate the accuracy for both the detection and the classification networks, four datasets are created, namely the classification train and test set and the detection train and test set. For the classification train set, the images were cropped such that the objects of interest occupied over 70 percent of the image. Examples for the umbilicus class are shown in Fig. 3(a). This enables the network to identify the differences between the two classes. Image augmentation operations like randomized pixel translation, pixel rotation, hue change, saturation change, and image mirroring were performed, making the classification train set six times the original size, consisting of 3600 wound and 2100 umbilicus images. This data was then randomly split into training and validation sets in 70:30 ratio. Further, the classification test set consists of 150 wound and 400 umbilicus images for testing the classification results. For training the Faster-RCNN, the detection training set consisted of images that would mimic the kind of images that would be taken by the tool camera on the robot. An example of the abdomen image is shown in Fig 3(b). The detection train set consists of 900 wound and 3000 umbilicus images whereas the detection test set consists of 150 images each for wound and umbilicus for testing the Faster-RCNN network.

2) Training Setup: Fig 3 shows the flowchart for training the Faster-RCNN for the umbilicus class. A similar flow is followed for the wound class as well.

For the feature extractor CNN we used a pre-trained network to implement transfer learning. AlexNet trained on ImageNet [12] was imported using MATLAB's Machine Learning Toolbox and the fully connected layer of AlexNet was modified to account for the two classes, wound and umbilicus. Using stochastic gradient descent method (SGDM) and an initial learning rate of 1e-4, the network was trained on the cropped images and converged in 10 epochs and 2450

iterations.

The trained CNN is now used as a feature extractor to simultaneously train the region proposal, and detection networks of Faster-RCNN. The positive overlap range was set to [0.6, 1], which ensured that a sample with the Intersection Over Union (IOU) of the Bounding Box (BB) greater than 0.6, would be classified as a positive training sample. Similarly, a negative overlap range of [0, 0.3] was used. Using SGDM, the network completed training the region proposal and detection networks in 1,232,000 iterations. The final detection results on the detection test data can be visualized in Figure 3(C).

Before determining the pixel coordinates of the objects in 2D image frame, the size of the BB was constrained to 300x300 pixels to remove false positives. This was experimentally determined by tracking the maximum possible pixel area for wounds and umbilicus, for the entire work space of the camera mounted on the end effector of the robot. Furthermore, only the BB with the highest score for umbilicus class was chosen as there can only be one umbilicus on the torso. For the wound class, the radius of the incircle of BB was given to the APF algorithm for wound avoidance.

For localization of the detected objects in the world frame, the *eye-in-hand* RGB-D camera was calibrated to the robot's end effector using the scheme described in [28]. Following this, the pixel coordinates (centroid of the BB) obtained from above were transformed to the robot's world frame to determine the object's location.

D. Autonomous Estimation of FAST locations

Using dimensions of the phantom obtained from the 3D torso map and visual landmarks (umbilicus), we estimate the FAST scan locations on the phantom in the robot's world frame using a scaling method inspired from [10]. In case of a free-hand scan, radiologists use the xiphoid (bottom tip of sternum) as a landmark to identify the 1st, 2nd, and 4th FAST positions (see Fig. 2(e)) on the patient.

Unfortunately, the location of the xiphoid is identified by touch and cannot be determined accurately by visual means. We instead use the umbilicus as a visual landmark in combination with the measured dimensions of the 3D phantom torso map to estimate the FAST locations. CT scan images of five anonymized atlas patients (see Section II-A.2.c) were used to identify a ratio (R_1) between the length of a person's torso (distance C in Fig. 2(d)) to the distance between the umbilicus and the 4th FAST position (distance A in Fig. 2(d)). A ratio (R_2) between C and distance of 3rd FAST scan region (distance B in 2(d)) from the umbilicus was calculated. Distance D represents the distance from the umbilicus to 1st, and 2nd FAST positions along y-axis. Another relation R_3 between C and D was calculated.

The width, length, and height of the phantom, along x, y, and z axes in Fig. 2(d) respectively, are estimated from the reconstructed 3D model. As the phantoms used were segmented torso models, their length is considered to be the length of the torso. Finally, using the above data, the system

calculates the FAST scan positions on the phantom from equations 1 and 2.

$$\begin{bmatrix} R_1 \\ R_2 \\ R_3 \end{bmatrix} = 1/C \begin{bmatrix} A \\ B \\ A - D \end{bmatrix} \quad \begin{bmatrix} X_1 \\ X_2 \\ X_3 \\ X_4 \end{bmatrix} = \begin{bmatrix} X_u + E \\ X_u - E \\ X_u \\ X_u \end{bmatrix} \quad (1)$$

$$\begin{bmatrix} Y_1 \\ Y_2 \\ Y_3 \\ Y_4 \end{bmatrix} = Y_u + \begin{bmatrix} R_3 C \\ R_3 C \\ -R_2 C \\ R_1 C \end{bmatrix} \quad \begin{bmatrix} Z_1 \\ Z_2 \\ Z_3 \\ Z_4 \end{bmatrix} = H \begin{bmatrix} 0.5 \\ 0.5 \\ 1 \\ 1 \end{bmatrix} \quad (2)$$

X_i, Y_i, Z_i for $i \epsilon 1, 2, 3, 4$ are coordinates of the respective FAST scan positions while X_u, Y_u, Z_u are the coordinates of the detected umbilicus in world frame. E is the width (see Fig. 2(d)) and H is the height of the patient.

The mean values from the atlas patients for R_1, R_2, and R_3 were found to be 0.29, 0.22, and 0.20, respectively.

E. Manual control - assisted tele-manipulation with haptic feedback

The remote radiologist uses a custom designed GUI (see Fig.2(g)) to monitor the system, and autonomously drive the robot to the FAST locations estimated above. The GUI provides control buttons and switches for the radiologist to select the FAST position to drive to, initiate autonomous 2D, 3D image generation, and toggling the tele-manipulation mode. To monitor the system, the GUI reports the amount of force exerted on the patient if the tool is in contact with the patient, and the status of the virtual fixture (VF) (described in Section II-E.1). The VF helps keep the probe in contact and apply a consistent contact force on the patient. The GUI also warns the radiologist if any wounds exist near a FAST location. The live views of tool and world cameras (shown in Fig.1 (b) and (c) respectively), and the Ultrasound (US) images were displayed using separate screens (see Fig. 2(f)).

Upon reaching the estimated FAST location, the robot is controlled using the haptic device based on live visual feedback from the *eye-in-hand* RGB tool camera. The world camera provides a bird's eye view of the scene for further clarity.

In our previous work, different control strategies were compared using tests that simulated the tele-manipulated ultrasonography system [29]. The rest of this section describes the tele-manipulation technique and force feedback algorithms when in contact with the patient from [29]. The force feedback algorithms before contact with the patient using APFs were developed as part of the current paper. We developed a *hybrid* control technique (described below)
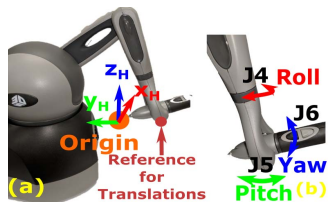

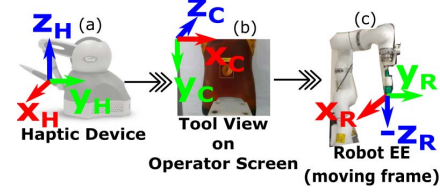Fig. 4: a) Haptic Device and b) Wrist Joints [29]

Fig. 5: Reference Frame Transformations [29]

using position control strategy for translation commands and rate control strategy for orientation commands. On evaluating the system's performance using multiple human operators, it was found that the *hybrid* control scheme was ideal for this application. The technique allowed the users to position the probe faster, as well as perform a FAST exam while smoothly sweeping the probe at more consistent angular velocities in each axis during reorientation of the probe during the scan.

The operator's commanded motions are given as increments to the robots initial Cartesian pose. In the case of this semi-autonomous system, this initial Cartesian pose is the pose of the robot after autonomously driving to a FAST location. The forward kinematics for translations of the haptic device are calculated with respect to the origin shown in Fig. 4(a) through the *Phantom Omni* driver [30]. To keep the system intuitive, the orientation commands are mapped using angular displacement of joints 4,5, and 6 from their center positions (Fig. 4(b)). The translations are thus controlled based on Eqn. (3).

$$\mathbf{X}_{R_{EE}} = H_w^T \mathbf{X}_{R_{EE_0}} + K_{P_1} \mathbf{X}_H, \quad (3)$$

where $\mathbf{X}_{R_{EE}} = [x\ y\ z\ 1]^T$ includes the new $x - y - z$ coordinates of the EE in current EE frame, $H_w^T \in \mathcal{R}^{4 \times 4}$ is the homogeneous transformation matrix from robot world frame to current EE frame, $\mathbf{X}_{R_{EE_0}} = [x_0\ y_0\ z_0\ 1]^T$ includes the initial positions of the EE in world frame, $K_{P_1} = \text{diag}[k_{p1}, k_{p2}, k_{p3}, 1] > 0$ are controller gains, and $\mathbf{X}_H = [x_H\ y_H\ z_H\ 1]^T$ are displacements of haptic device from origin shown in Fig. 4.a.

The rate of change of each orientation is directly proportional to the deviation of each axis (J4, J5, J6 as shown in Fig. 4.b) from its mean position. A dead-zone of -35 degrees to +35 degrees in roll and yaw, and -20 degrees to +20 degrees in pitch axis was created to avoid accidental change in EE orientation. The orientation control scheme can be mathematically represented as:

$$\dot{\theta}_{R_{EE}} = K_\theta \theta_H, \quad (4)$$

where, $\dot{\theta}_{R_{EE}} \in \mathcal{R}^3$ includes the roll, pitch, and yaw rates of the EE in *world* frame, $K_\theta = \text{diag}[k_{\theta 1}, k_{\theta 2}, k_{\theta 3}] > 0$ are controller gains, and $\theta_H \in \mathcal{R}^3$ are wrist joint angles of the haptic device (Fig.4.b). The $\dot{\theta}_{R_{EE}}$ is integrated at a constant loop rate of 1000 Hz to obtain $\theta_{\mathbf{R_{EE}}}$. The Cartesian pose formed from $\mathbf{x}_{R_{EE}}$ and $\theta_{R_{EE}}$ is first transformed to world frame, then converted to robot's joint space using inverse kinematics, before feeding them to the robot.

653

1) Force feedback algorithms when in contact with the patient: Haptic force feedback for the normal forces, i.e. forces acting along the z_R - axis (see Fig. 5) was provided to the radiologist through the haptic device. Only normal forces were used because the US image is not affected by the lateral forces applied by the scanning area of the probe. A gravity compensation scheme as described in [31] was used to neutralize the weight of the tool on the force sensor. For safety, a hard virtual fixture (VF) was incorporated to shutdown the system if the force threshold of 12N was breached. This threshold is based on previous studies [32]. In [29], a soft VF was developed to "lock" the EE's translatory motion while still allowing orientation control to the user if the probe was exerting a normal force greater than a threshold value. The VF prevents any translation except in the $+z_R$ - axis shown in Fig. 5. This enables the user to continue scanning and deactivate the virtual fixture by moving the probe away from the patient. It was observed that the interaction force may increase during reorientation while the probe is in contact with the patient. The VF helps limit the forces applied on the patient while keeping the probe stable during sweeping scanning motions. The force threshold for the soft virtual fixture was set to 3N determined experimentally based on US image quality.

On evaluation of the virtual fixture in [29], it was found that the users were able to maintain contact of the probe for around 82 % of the duration of the scan with the VF as compared to only 40 % in the case without the VF. The standard deviation of forces was also 150 % smaller in the case with VF indicating higher consistency in contact forces.

2) Force feedback algorithms before contact with the patient: APFs emerging from the center of the wound in the direction of the EE are generated in the shape of a hemisphere with a radius 2 cm greater than the radius of the identified gunshot wound as shown in Fig. 2 (f). The forces exerted on the EE by the APFs are reflected to the user as haptic feedback to help guide the probe away from the wounds.

III. EXPERIMENTS AND RESULTS

A. Umbilicus and Wound Classification and Localization

a) Classification and Detection: In this experiment, we evaluate the system's ability to accurately classify umbilicus and wounds on the classification test set and then calculate the detection accuracy on the detection test set (described in section II-C.1). From Table I, we can infer that the sensitivity (recall or true positive rate) and specificity (True Negative rate) of the feature extractor CNN for both the classes is above 94 %. This can therefore be used as a base network to train the Faster-RCNN model. The overall accuracy for the classifier was 97.9 %.

Class Name	Sensitivity (Recall)	Specificity	Accuracy
Wounds	96.08 %	94.78 %	97.9 %
Umbilicus	98.3 %	98.14 %	

TABLE I: Evaluation results for Feature Extractor CNN.

Once the Faster R-CNN model was trained, the Mean Average Precision (mAP) was calculated on the detection test set at an Intersection Over Union (IOU) of 0.5. This metric was followed as per the guidelines established in PASCAL VOC 2007 [18]. The mAP for the two classes at an IOU of 0.5 is 0.51 and 0.66 for umbilicus and wounds, respectively.

b) Localization: The robotic system generates a warning for the radiologist if a wound is detected near a FAST exam location. Moreover, the FAST exam points are estimated with respect to the umbilicus. Hence, it is important to accurately determine the position of these objects in the robot's frame. In this experiment, we take readings for objects of each of the three classes placed at 5 random positions and angles on the wound phantom (described in Section II-A.2.a) to make a total of 15 tests. The ground truth values for each, umbilicus and wounds were estimated by touching the actual center of the object using a pointed tool attached to the robot and performing forward kinematics to determine the object locations in the robot world frame.

Class Name	Mean Error(mm)	Standard Deviation(mm)
Umbilicus	8.77	1.50
Wounds	9.30	0.79

TABLE II: Localization errors for umbilici and wounds.

The results for average error (Euclidean distance) and standard deviation for each class is reported in II. The average localization error for both the classes combined was found to be 0.947 cm $\pm$ 0.179 cm.

B. Estimating FAST Scan Locations

To evaluate the accuracy of the estimated FAST exam points, the five localization phantoms were scanned with our robotic system and the four FAST scan positions were estimated using methods described in Section II-D. The actual positions are the centroids of the FAST regions on the localization phantoms (Section II-A.2.b). FAST scan positions 1 and 2 are placed symmetrically on the body. Hence, their accuracies are reported together.

Since the final step of the semi-autonomous FAST exam is tele-manipulated, the estimated locations need not be highly accurate and only need to be within the workspace of the haptic device and Field Of View (FOV) of the camera. Table III shows the mean error (Euclidean distance) and standard deviation between the estimated and actual positions of the four FAST exam points in the robot's base frame. The third column of Table III shows whether the estimated point was within this marked region. The average position accuracy of the system was 2.2 cm $\pm$ 1.88 cm. All the estimated points were found to be within the scanning region marked by the expert radiologist for each FAST scan location. The largest error for any location over the five test phantoms was 7.1 cm, which was well within the workspace of the slave robot and FOV of the RGB-D camera. Thus, all the FAST points were within tele-manipulable distance from the estimated initialization positions.

	Mean Error (cm)	Standard Deviation (cm)	Inside Marked Region	Inside Workspace
Pt 1, 2	1.66	1.46	Yes	Yes
Pt 3	1.64	0.87	Yes	Yes
Pt 4	3.34	3.31	Yes	Yes

TABLE III: Estimated and actual FAST scan positions.

C. Artificial Potential Fields (APFs)

In this experiment, the subjects were asked to tele-manipulate the robot from the start to the end point while trying to avoid a wound of 4cm radius in the path. The resulting trajectories for 4 different subjects from a pilot study involving 8 human subjects can be seen in Fig. 6. The color heat map defines the amount of resultant force being exerted by the APF. From the 3 colored trajectories (with APF), it can be seen that the APF was able to guide the probe around the hemisphere and successfully avoid the wound. The trajectory in black (without APF) was able to avoid the wound but still allowed the subject to maneuver the probe very close to it. Overall, the subjects maintained an average distance of 5.09 cm in the presence of the APFs with a minimum distance of 4.575 cm to the center of the wound. However, in the absence of the APFs, the subjects maintained an average distance of 3.3 cm with a minimum distance of 1.8 cm to the center of the wound, and caused one collision with the wound. A *t-tail* test for the two groups of distances between the EE and the wound, resulted in a *p-value* of 0.0117 proving that the data is statistically conclusive.

D. Trauma Assessment of FAST Phantom

In this experiment, all the sub-systems, namely, autonomous 2D and 3D scanning, 2D image analysis, 3D Map generation, autonomous estimation of FAST locations, autonomous positioning of US probe at FAST exam positions, and assisted tele-manipulation with haptic feedback were integrated together to perform a complete trauma assessment on a FAST Ultrasound phantom. The test started off with the 3D reconstruction of the FAST phantom as shown in Fig. 7. The image of an umbilicus was pasted on the phantom at an approximate anatomically correct location recommended by the radiologist. The FAST locations were thus estimated using the size of the phantom obtained from the reconstruction of the phantom, and localization of the umbilicus. The estimation error was found to be 2.7 cm, 0.56 cm, and 0.43 cm for positions 3, 4, and 1 and 2, respectively. These were well within the slave's workspace and hence, easily tele-manipulable from the autonomously initialized
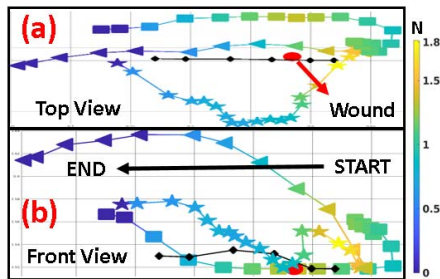

Fig. 7: 3D point cloud model of the phantom.

position of the US probe. Wound detection and classification was not tested in this experiment to avoid any permanent damage to the phantom from the artificial colorings used to simulate the object.

A 34-minute long training about the robotic system including a practice session for tele-manipulation was given to the expert radiologist. Following the training, a complete FAST scan of the US phantom was conducted. The trauma assessment using the robotic system was completed in 16 minutes and 3 seconds as compared to freehand manual scanning, which was completed in 4 minutes and 13 seconds. Even though robotic system is much slower than a free-hand scan, speed is of little significance considering that this exam is conducted remotely while transporting the trauma patient to a hospital saving invaluable time which otherwise would have been wasted. The maximum normal force recorded during the complete procedure was 6.3 N which is well below our safety limit of 12 N. The hard virtual fixture was thus never activated during the actual examination. A blinded radiologist graded the nine ultrasound screenshot images that best displayed the extent of the hemorrhages in the FAST phantom for each the robotic system and a free hand scan. The images, which were randomized to avoid any bias, were graded on a Likert scale (1-low to 5-high). The images from the robotic system received a score of 44 of possible 45 as compared to a 41 for the images from the free hand scan. Standard deviations of 0.33 and 0.53 were obtained for the robotic system and hand scan images, respectively. The two-tailed *p-value* for the scores was found to be 0.1284, which shows that the difference in the scores requires more data to be statistically significant. Nonetheless, the results show that the US images produced by the robotic system are at least on par with those obtained by an expert radiologist by hand. Representative images of regions of interest from each of the four FAST points along with scores received (marked in yellow) are shown in Fig. 8.
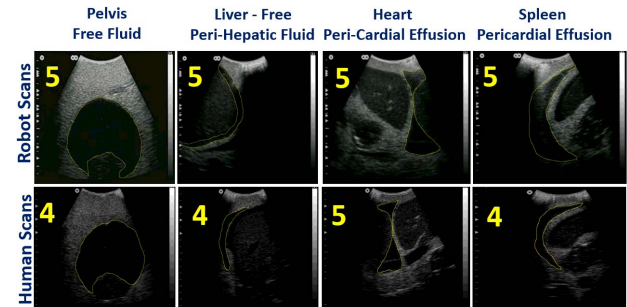

Fig. 6: Trajectories for APF test. a) Top View. b) Front View.

Fig. 8: Ultrasound scan comparison with individual image quality scores (1-low, 5-high).

655

IV. Discussion & Conclusion

This paper reported the development of a semi-autonomous robotic system for trauma assessment. The system accurately calculates FAST scan positions of an ultrasound phantom using the measured phantom size and position of the umbilicus. The system also accurately detects and localizes gunshot wounds on the patient's torso to help the physician avoid them during tele-manipulation using Artificial Potential Fields. The assisted tele-manipulation strategy was found to be successful when analyzed using multiple human operators. A radiologist successfully completed a remote FAST scan of the phantom using the system. These results combined with the high quality of the obtained ultrasound images show early feasibility of the system and demonstrate the potential to virtually bring a physician to an ambulance for pre-hospital trauma assessment of a patient. As part of future work, impedance control may be used to apply more consistent forces during the US scan. The wound detection can be extended to more types of trauma injuries. The ability to detect more anatomical landmarks using the 3D images can be added to enable the calculation of torso length of an actual patient.

References

[1] N. C. for injury Prevention & Control, "National vital statistics system." [Online]. Available: https://www.cdc.gov/injury/wisqars/LeadingCauses.html

[2] D. E. Clark, "RA cowley, the golden hour, the momentary pause, and the third space," *The American Surgeon*, vol. 83, no. 12, pp. 1401–1406, 2017.

[3] R. Edlich, J. R. Wish, L. D. Britt, and W. B. Long III, "An organized approach to trauma care: legacy of r adams cowley," *Journal of long-term effects of medical implants*, vol. 14, no. 6, 2004.

[4] J. S. Davis, S. S. Satahoo, F. K. Butler, H. Dermer, D. Naranjo, K. Julien, R. M. Van Haren, N. Namias, L. H. Blackbourne, and C. I. Schulman, "An analysis of prehospital deaths: Who can we save?" *Journal of Trauma and Acute Care Surgery*, vol. 77, no. 2, pp. 213–218, 2014.

[5] L. J. Wherrett, B. R. Boulanger, B. A. McLellan, F. D. Brenneman, S. B. Rizoli, J. Culhane, and P. Hamilton, "Hypotension after blunt abdominal trauma: the role of emergent abdominal sonography in surgical triage." *The Journal of trauma*, vol. 41, no. 5, pp. 815–20, nov 1996.

[6] J. Montoya, S. P. Stawicki, D. C. Evans, D. P. Bahner, S. Sparks, R. P. Sharpe, and J. Cipolla, "From FAST to e-FAST: an overview of the evolution of ultrasound-based traumatic injury assessment," *European journal of trauma and emergency surgery*, vol. 42, no. 2, pp. 119–126, 2016.

[7] S. M. Mazur, A. Pearce, S. Alfred, A. Goudie, and P. Sharley, "The F.A.S.T.E.R. trial. Focused assessment by sonography in trauma during emergency retrieval: a feasibility study," *Injury*, vol. 39, no. 5, pp. 512–518, May 2008.

[8] K. Ito, S. Sugano, and H. Iwata, "Wearable echography robot for trauma patient," in *2010 IEEE/RSJ International Conference on Intelligent Robots and Systems*, Oct 2010, pp. 4794–4799.

[9] H. T. Sen, A. Cheng, K. Ding, E. Boctor, J. Wong, I. Iordachita, and P. Kazanzides, "Cooperative control with ultrasound guidance for radiation therapy," *Frontiers in Robotics and AI*, vol. 3, p. 49, August 2016.

[10] A. S. B. Mustafa, T. Ishii, Y. Matsunaga, R. Nakayear, H. Ishii, K. Ogawa, A. Saito, M. Sugawara, K. Niki, and A. Takanishi, "Development of robotic system for autonomous liver screening using ultrasound scanning device," in *2013 IEEE International Conference on Robotics and Biomimetics (ROBIO)*. IEEE, 2013-12, pp. 804–809.

[11] D. R. Swerdlow, K. Cleary, E. Wilson, B. Azizi-Koutenaei, and R. Monfaredi, "Robotic ArmAssisted Sonography: Review of Technical Developments and Potential Clinical Applications," *American Journal of Roentgenology*, vol. 208, no. 4, pp. 733–738, Feb 2017.

[12] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks," *Advances In Neural Information Processing Systems*, pp. 1–9, 2012.

[13] J. Deng, W. Dong, R. Socher, L. Li, and and, "Imagenet: A large-scale hierarchical image database," in *2009 IEEE Conference on Computer Vision and Pattern Recognition*, June 2009, pp. 248–255.

[14] E. Nasr-Esfahani, S. Samavi, N. Karimi, S. Soroushmehr, M. Jafari, K. Ward, and K. Najarian, "Melanoma detection by analysis of clinical images using convolutional neural network," *38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pp. 1373–1376, Aug. 2016.

[15] F. Li, C. Wang, X. Liu, Y. Peng, and S. Jin, "A composite model of wound segmentation based on traditional methods and deep neural networks," *Computational Intelligence and Neuroscience*, vol. 2018, no. 1, 2018.

[16] P. Sermanet and D. Eigen, "OverFeat : Integrated Recognition , Localization and Detection using Convolutional Networks arXiv : 1312 . 6229v4 [cs . CV] 24 Feb 2014."

[17] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, no. 6, pp. 1137–1149, 2017.

[18] J. M. Winn, M. Everingham, L. V. Gool, C. K. I. Williams, and J. Winn, "The Pascal Visual Object Classes (VOC) challenge The P ASCAL Visual Object Classes (VOC) Challenge," no. June, 2010.

[19] T.-Y. Lin, M. Maire, S. Belongie, L. Bourdev, R. Girshick, J. Hays, P. Perona, D. Ramanan, C. L. Zitnick, and P. Dollár, "Microsoft COCO: Common Objects in Context," may 2014. [Online]. Available: https://arxiv.org/abs/1405.0312

[20] M. Quigley, K. Conley, B. Gerkey, J. Faust, T. Foote, J. Leibs, R. Wheeler, and A. Y. Ng, "ROS: an open-source robot operating system," in *ICRA workshop on open source software*, vol. 3. Kobe, Japan, 2009, p. 5.

[21] R. Smits, "KDL: Kinematics and Dynamics Library," http://www.orocos.org/kdl.

[22] C. Hennersperger, B. Fuerst, S. Virga, O. Zettinig, B. Frisch, T. Neff, and N. Navab, "Towards MRI-based autonomous robotic US acquisitions: A first feasibility study," *IEEE Transactions on Medical Imaging*, vol. 36, no. 2, pp. 538–548, 2017-02.

[23] P. J. Besl and N. D. McKay, "A method for registration of 3-d shapes," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 14, no. 2, pp. 239–256, Feb 1992.

[24] D. Holz, A. E. Ichim, F. Tombari, R. B. Rusu, and S. Behnke, "Registration with the point cloud library: A modular framework for aligning in 3-d," *IEEE Robotics Automation Magazine*, vol. 22, no. 4, pp. 110–124, Dec 2015.

[25] F. Pomerleau, F. Colas, R. Siegwart, and S. Magnenat, "Comparing ICP Variants on Real-World Data Sets," *Autonomous Robots*, vol. 34, no. 3, pp. 133–148, Feb. 2013.

[26] R. B. Rusu and S. Cousins, "3d is here: Point cloud library (pcl)," in *2011 IEEE International Conference on Robotics and Automation*, May 2011, pp. 1–4.

[27] J. R. R. Uijlings, K. E. A. Van De Sande, T. Gevers, and A. W. M. Smeulders, "Selective Search for Object Recognition," *International Journal of Computer Vision*, vol. 104, no. 2, pp. 154–171, 2012.

[28] R. Y. Tsai and R. K. Lenz, "A new technique for fully autonomous and efficient 3d robotics hand/eye calibration," *IEEE Transactions on Robotics and Automation*, vol. 5, no. 3, pp. 345–358, June 1989.

[29] B. Mathur, A. Topiwala, H. Saeidi, T. Fleiter, and A. Krieger, "Evaluation of Control Strategies for a Tele-manipulated Robotic System for Remote Trauma Assessment," *Conference on Control and Its Applications, SIAM*, 2019.

[30] B. Mathur, "Geomagic_Touch_ROS_Drivers," https://github.com/bharatm11/Geomagic$_Touch_ROS_Drivers/$.

[31] L. Richter, R. Bruder, and A. Schweikard, "Calibration of force/torque and acceleration for an independent safety layer in medical robotic systems," *Cureus*, vol. 4, no. 9, 2012.

[32] M. A. L. Bell, H. T. Sen, I. Iordachita, and P. Kazanzides, "Force-controlled ultrasound robot for consistent tissue pre-loading: Implications for acoustic radiation force elasticity imaging," in *5th IEEE RAS/EMBS International Conference on Biomedical Robotics and Biomechatronics*. IEEE, 2014-08, pp. 259–264.