Proceedings of The 2024 IEEE International
Conference on Real-time Computing and Robotics
June 24 - 28, 2024

2024 IEEE International Conference on Real-time Computing and Robotics (RCAR) | 979-8-3503-7260-1/24/$31.00 ©2024 IEEE | DOI: 10.1109/RCAR61438.2024.10670924

A Dynamical System Approach to Robotic Ultrasound Imaging: Towards Intrinsically Stable Robotic Sonography

Wanli Liuchen, Anqing Duan, Zirui Song, Maria Victorova, and David Navarro-Alarcon

***Abstract*— In this research, we introduce a framework that leverages the deep learning capabilities of the Residual Network (ResNet) architecture to improve the efficiency of robotic systems in ultrasound-guided body scanning. It helps accurately spot abnormalities in tissues and understand the detailed aspects of ultrasound images. Subsequently, the framework capitalizes on the strengths of time-invariant dynamical systems (DS) to refine the efficiency of robotic manipulation. Contrasting with conventional approaches necessitating manual probe maneuvering across the patient's body while maintaining consistent contact force, our methodology permits the robotic arm to fulfill multiple objectives concurrently, such as sustaining uniform contact pressure and achieving exact positioning. By preemptively adjusting the robot's dynamics to ensure a stable interaction with the patient's surface, our system addresses minor bodily movements, maintaining constant contact and ensuring a comprehensive and precise scanning process.**

Fig. 1: A sonographer is doing an abdominal scan.

I. INTRODUCTION

In recent decades, there has been significant attention on healthcare and medical assistive robots due to their escalating demand in clinical settings [1]. Among these robots, those equipped with ultrasound-guided navigation for physical body examination have found extensive applications [2]. Unlike conventional X-rays, ultrasound possesses several advantages, including the absence of ionizing radiation and affordability, making it a preferred choice in a plethora of medical fields such as cardiology, urology, and gynecology [3].

The paramount importance of body tissue examination lies in its ability to provide timely detection and intervention for various medical conditions. As the medical field persistently strives to minimize exposure to ionizing radiation from X-ray scans, the technique of 3D tissue reconstruction via ultrasound has gained significant attention. Traditionally, this process involves a human operator meticulously scanning the patient by guiding an ultrasound probe over the target area as illustrated in fig. 1. Achieving optimal image quality requires real-time probe adjustments to highlight structures or organs [4].

There has been an increasing number of researchers developing applications for robotic technologies in 3D ultrasound (US) imaging over the past few decades [5], [6]. These applications, by employing robotic mechanisms to drive the scanning probes, enable the system to track and scan with precision, thereby enhancing the accuracy of volumetric construction. However, current robotic-based diagnostic approaches predominantly encounter two critical challenges: the issue of contact between human tissue and the robot, and the technology related to positioning [6].

This paper presents a framework that integrates a Dynamical System (DS) controller with Convolutional Neural Network (CNN)-based methods, offering a more robust and comprehensive control strategy to address these challenges. Our framework's efficacy and innovative aspects are substantiated through simulation and empirical results. The Dynamical System (DS) method has been an important tool in robotic controls, especially in tasks demanding high adaptability and real-time responses. The DS approach offers an intuitive means to describe and predict the behavior of complex robotic systems in dynamic environments [7]. By employing DS methods in combination with impedance control, our objective is to validate the applicability of the ideal model in real-world scenarios, specifically by verifying the system's tracking performance, mechanical control capabilities, and scanning and localization precision. This, in turn, enhances the robot's adaptability for body tissue examination.

Furthermore, the majority of existing methodologies tend to focus primarily on stiffness control at contact points [8]. Our research diverges from these conventional methods by emphasizing not only stiffness but also regulating the force applied. This dual focus, we believe, offers a more holistic control strategy, ensuring both the stability and accuracy of the robotic arm during its operation. By meticulously integrating these advanced control strategies, striving for enhanced precision and efficiency in robotic-assisted tissue examinations in clinical environments.

W. Liuchen, A. Duan, Z. Song, and D. Navarro-Alarcon are with the Faculty of Engineering, The Hong Kong Polytechnic University, KLN, Hong Kong.

M. Victorova is with Bloom Standard, Shatin, NT. Hong Kong.

II. METHODOLOGY

A. Object Detection and Localization

Deep learning techniques, particularly those tailored for object detection and localization, have shown significant promise in recent years. The methodology chosen for this project is inspired by and builds upon the foundational work presented in [9]. This selection was made after careful consideration of alternative models, such as VGG11 and DenseNet121, which are also recognized as potential backbone models. However, it was found that their performance efficiency on ultrasound images was comparatively lower [10].

Central to our methodology is the Residual Network (ResNet) architecture, a renowned convolutional neural network (CNN) [11]. Serving as the system's backbone, this network establishes the primary structure upon which further components are integrated, as detailed in fig. 2.

Our model introduces several modifications to the canonical ResNet18, tailored specifically for our objective of object detection and localization. Instead of the traditional final fully connected layer designed for classification scores, we substituted it with a trio of deconvolutional layers. This adjustment is crucial for our goal to accurately detect and localize objects, as it shifts the model's output from class probabilities to spatial maps. The decision to implement these deconvolutional layers, inspired by insights from [12], includes integrating batch normalization to enhance network stability and performance. Additionally, the inclusion of the rectified linear unit (ReLU) activation function, based on recommendations from [13]:

$$f(x) = \max(0, x)$$

where x is the input signal. The ReLU introduces necessary non-linearity, boosting the model's learning capability.

The deconvolutional layers, coupled with a concluding 1×1 convolutional layer, play a vital role in generating heatmaps that highlight potential object locations. The peak intensity within these heatmaps precisely marks the pixel most likely to represent the object's location, serving as a guide for pinpointing its central position. We calculate the loss between the actual values and the predicted heatmap using the mean squared error (MSE), refining the model's precision in pinpointing object locations. The loss function is defined as:

$$\mathcal{L}_{\text{mse}} = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2$$

where n represents the number of samples, y_i denotes the true value for sample i, and $\hat{y}_i$ is the predicted value for sample i.

B. Robot Dynamics

In this context, the variable x signifies the Cartesian position, and an integral curve of $f(x, \dot{x})$ delineates the anticipated motion trajectory of the robot in the absence of external disturbances. The second-order dynamical system characterizing $f(x, \dot{x})$ is:

$$f(x, \dot{x}) = -A(\dot{x}) - B(x - x^t) \tag{1}$$

where $A \in \mathbb{R}^3$ represents the damping coefficient corresponding to velocity, capturing the effects of velocity dissipation. The matrix $B \in \mathbb{R}^{3\times3}$ depicts the potential energy matrix across distinct coordinate axes. x^t stands for the target state of the robot. Consequently, $(x - x^t)$ highlights the discrepancy between the robot's present and desired states. For a robotic manipulator with N-degrees-of-freedom operating in a three-dimensional Cartesian space, its dynamics can be modeled as:

$$M(x)\ddot{x} + C(x, \dot{x})\dot{x} + g(x) = F_c - F_m \tag{2}$$

where $x \in \mathbb{R}^3$ represents the robot's position, $M(x) \in \mathbb{R}^{3\times3}$ is the mass matrix, $C(x, \dot{x})\dot{x} \in \mathbb{R}^3$ denotes the centrifugal and Coriolis forces, and $g(x) \in \mathbb{R}^3$ accounts for the gravitational forces. $F_c \in \mathbb{R}^3$ and $F_m \in \mathbb{R}^3$ correspond to the control forces and the externally measured forces, respectively. Specifically, F_c is crucial for enabling the robot to follow a predefined velocity trajectory $\dot{x}_d \in \mathbb{R}^3$.

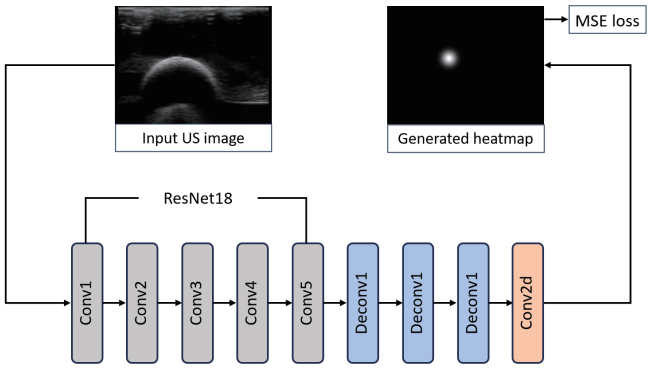

Fig. 2: We utilize a fully connected network based on the ResNet architecture with a deconvolutional head. The loss is computed using the mean squared error between the target and predicted heatmaps. The final pose is extracted from the point of maximum intensity on the predicted heatmap[9].

C. Controller Design

The control force F_c is derived from the DS-impedance controller as described in [14] and [15]:

$$F_c = D(x, \dot{x})f(x, \dot{x}) \tag{3}$$

During the manipulation process, direct control is applied to the end-effector. Consequently, the DS is utilized solely for the translation of the end-effector. $D(x, \dot{x}) \in \mathbb{R}^{3\times3}$ serves as a modulation function, adjusting the nominal DS in alignment with the contact surface, contingent upon the robot's current state [16]. To ensure stable tracking of the target, the first eigenvector $d_1 \in \mathbb{R}^+$ of matrix $D(x, \dot{x}) \in \mathbb{R}^{3\times3}$ aligns with the desired dynamical motion x_d. The varying damping term dissipates selectively in directions orthogonal to the desired direction of motion given by $f(x, \dot{x})$. The state-varying damping matrix, denoted as $\mathsf{D}(x, \dot{x})$, is formulated

as [16]:

$$\mathsf{D}(x, \dot{x}) = Q\mathcal{L}Q^T \tag{4}$$

where Q is an orthonormal basis in $\mathbb{R}^d$ represented by the matrix $Q = \begin{bmatrix} q_1 & \dots & q_d \end{bmatrix}$. Each vector q_i for $i \in \{1, \dots, d\}$ forms a component of this basis.

The elements of the matrix $\mathcal{L}$, represented as $\mathbf{\Delta}_{ij}(x, \dot{x})$ for $i, j \in \{1, \dots, d\}$, are crucial in controlling the robot's movement. They allow for adjustments in both tangential and normal directions relative to the contact surface. Specifically, we set $\mathbf{\Delta}_{1j}(x, \dot{x}) = 0$ to eliminate the robot's acceleration perpendicular to the surface. To steer the robot in the q_i direction using the nominal DS, we configure $\mathbf{\Delta}_{ii}(x, \dot{x}) = 1$ and $\mathbf{\Delta}_{ij}(x, \dot{x}) = 0$ for all $i, j \in \{1, \dots, d\}$ where $i \neq j$.

The 'transition region' is a specific area near the contact surface where the modulation effect is limited. This region includes all points where $0 < \mathcal{G}(x) \leq \rho$, with ρ being a positive real number. As the robot moves beyond this region, the modulation effect decreases exponentially with increasing distance from the surface. The modulation of the DS dynamics is specified by [16]:

$$\mathbf{\Delta}_{ij}(x, \dot{x}) = \begin{cases} \Delta_{ij}(x, \dot{x}) & \text{if } \mathcal{G}(x) \leq \rho \\ \left(\Delta_{ij}(x, \dot{x}) - 1\right) e^{\frac{\rho - \mathcal{G}(x)}{\sigma}} + 1 & \text{if } i = j, \rho < \mathcal{G}(x) \\ \Delta_{ij}(x, \dot{x}) e^{\frac{\rho - \mathcal{G}(x)}{\sigma}} & \text{if } i \neq j, \rho < \mathcal{G}(x) \end{cases} \tag{5}$$

Here, σ defines how rapidly the modulation effect decreases in the free motion area. When $\rho < \mathcal{G}(x)$, meaning the robot is considerably distant from the contact surface, $\mathcal{L}$ becomes the identity matrix. This indicates that the nominal dynamical system exclusively dictates the robot's motion.

Due to the fact that $f(x, \dot{x})$ is an eigenvector of $D(x)$, we can rewrite F_c as:

$$\ddot{x} = q_1 \dot{x}_d - D(x)\dot{x} \tag{6}$$

To achieve the desired contact motion and force control post-contact using a single dynamical system (DS), we express $\dot{x}_d$ as follows :

$$\dot{x}_d = f(x, \dot{x}) + f_n(x, \dot{x}) \tag{7}$$

where $f_n(x, \dot{x})$ is the modulation term applied perpendicularly to the surface to manage post-contact motion, described in [14] as:

$$f_n(x, \dot{x}) = \frac{F_d(x, \dot{x})}{d_1} n(x) \tag{8}$$

We assume the contact surface is impenetrable, with the normal vector $n(x)$ and the distance to the surface $\mathcal{G}(x)$ known at every point in space. These parameters can be estimated using different regression-based learning methods [17], [18]. The desired force profile, $F_d(x) \in [0, F_{\max}]$, where $F_{\max} > 0$, specifies the target force range and depends on the state.

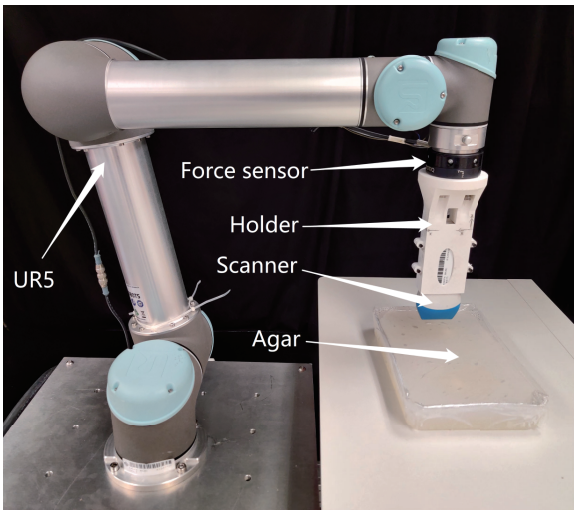

Fig. 3: Experiment setup

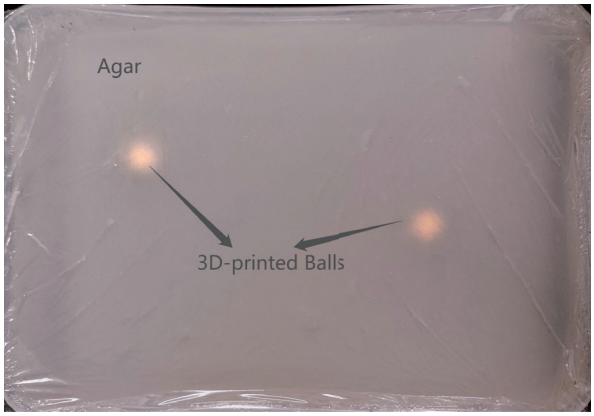

Fig. 4: Agarose gel is used in ultrasound examinations to simulate human tissue. 3D-printed spheres are incorporated into the gel to simulate abnormal pathological conditions.

III. EXPERIMENTAL EVALUATION

A. Experiment Setup

In our study, we emulated human tissue conditions using agar gel depicted in fig. 4, known for its similarity to human tissue in ultrasound examinations. Within this gelatinous medium, we inserted objects that mimic abnormal pathological conditions, thereby generating a more accurate representation of real-world medical scenarios.

To create our training and testing datasets, we performed a series of ultrasound scans on the simulation model. The data gathered from these scans consisted of 320 training and 40 testing images, providing a substantial foundation for the model's learning process. The scanned images and their corresponding heatmaps, which signify the model's object detection, are presented in fig. 5. We employed deep learning techniques to train our model on the generated dataset, iterating until we achieved an overall accuracy of 80% We found this result to be satisfactory for our initial experiments, given the complexity of the task at hand. Our system is engineered for real-time operations, allowing for both data

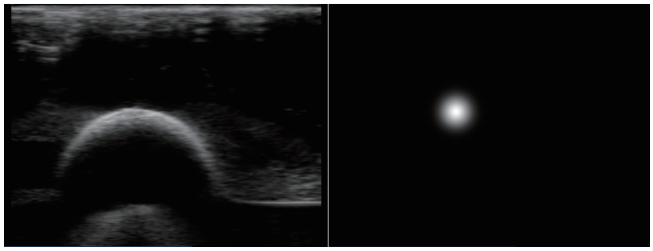

Fig. 5: The original scanned image is on the left, and the processed heatmap is on the right

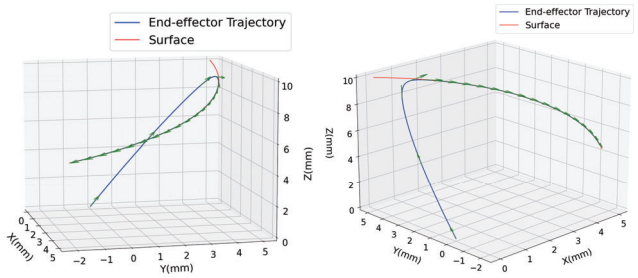

Fig. 6: The control involves both the nominal dynamical system and modulation.

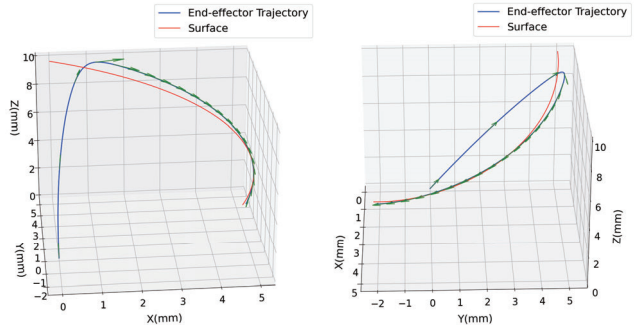

Fig. 7: Only the nominal dynamical system is used for control, without any modulation.

input from and output to a local notebook computer. This capability facilitates live experiments and enables us to adjust our model based on immediate feedback, thereby optimizing our testing process and enhancing the model's performance.

We focus on a surface scanning task, evaluating the proposed framework on a real robotic arm platform, specifically a 6-degree-of-freedom robotic arm (Universal Robots UR5) as shown in fig. 3. The robot operates at approximately 100 Hz, controlling the end-effector position in the horizontal plane. The robotic arm translates the output of the dynamical system (DS) into joint states automatically. Additionally, the robot is equipped with a 6-axis force-torque sensor (FT 300 Force Torque Sensor), which records force data but is not used in the control loop. The nominal second-order dynamical system of the DS is defined as:

$$f(x, \dot{x}) = -20\dot{x} - 100I(x - x^t) \tag{9}$$

where I denotes the identity matrix of the appropriate dimension. These parameters were selected to ensure the robot enters the transition region effectively.

B. Simulation

In the presented simulation, a dynamic system is steered towards a moving target using a control mechanism. The system's motion is primarily governed by a nominal dynamical system, which is augmented with a controller to ensure precise tracking of the moving target. This controller varies with the state and velocity of the system and is constructed based on an orthonormal basis and a modulation matrix. The modulation matrix, in particular, plays a pivotal role in directing the robot's motion both tangentially and normally concerning a contact surface, thereby offering a nuanced control over its trajectory. To further quantify the accuracy of this control mechanism, we utilized the Mean Squared Error (MSE) to calculate the deviation between the trajectory of the robot's end effector and the moving target's trajectory over the latter 50% of the data points, providing a targeted measure of trajectory error in the latter stages.

The simulation results distinctly illustrate the controller's effectiveness. The system achieves an MSE of 0.0 ± 0.0301, as detailed in table I. This precision, depicted in fig. 6, demonstrates the controller's ability to adeptly follow the target's intricate movement patterns with minimal lag and overshoot, thereby showcasing its rapid adaptability to changing conditions and guiding the system in real-time with remarkable accuracy.

Conversely, a subsequent simulation that proceeded without the additional modulation—thus operating solely on the nominal dynamical system—registered a significantly higher MSE of 0.1 ± 0.042, as recorded in table I. The corresponding results, illustrated in fig. 7, indicate that while the system still directed itself towards the target, it lacked the precision and adaptability observed with additional modulation. This led to pronounced tracking errors and a marked decrease in the system's capacity to adjust to the dynamic changes of the moving target's trajectory.

C. Performance Evaluation

To rigorously evaluate our method's efficacy, we focused on the force tracking error by comparing the measured and desired force values. The chosen evaluation metric, the Root Mean Square Error (RMSE):

$$e = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}$$

Experimental set-up: Tracking Simulation		
	w/ modulation	w/o modulation
Initial Position	$[5 \pm 0.3,$ $5 \pm 0.2,$ $10 \pm 0.5]$	$[5 \pm 0.2,$ $5 \pm 0.3,$ $10 \pm 0.4]$
Mean Squared Error	0.0 ± 0.0301	0.1 ± 0.042

TABLE I: The values of MSE with and without additional modulation were compared when using a moving target with a slight adjustment of the initial position

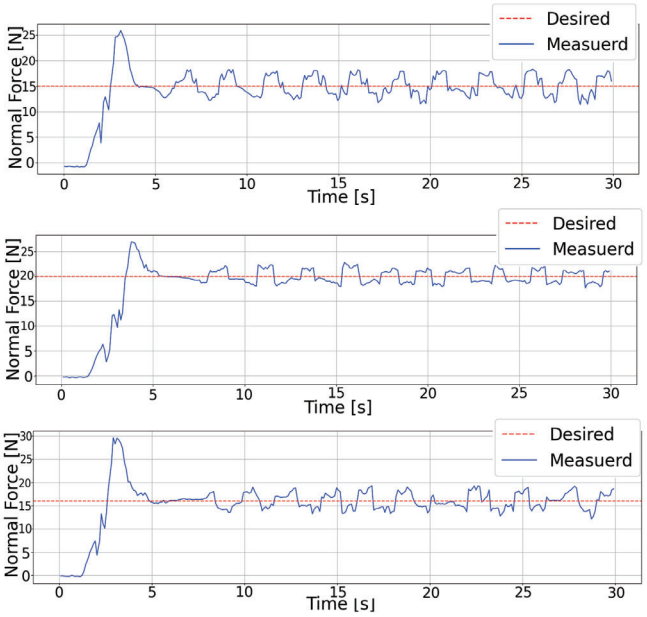

Fig. 8: We set target forces of 15N, 16N, and 20N to test the mechanical tracking performance of the dynamic system after our modulation.

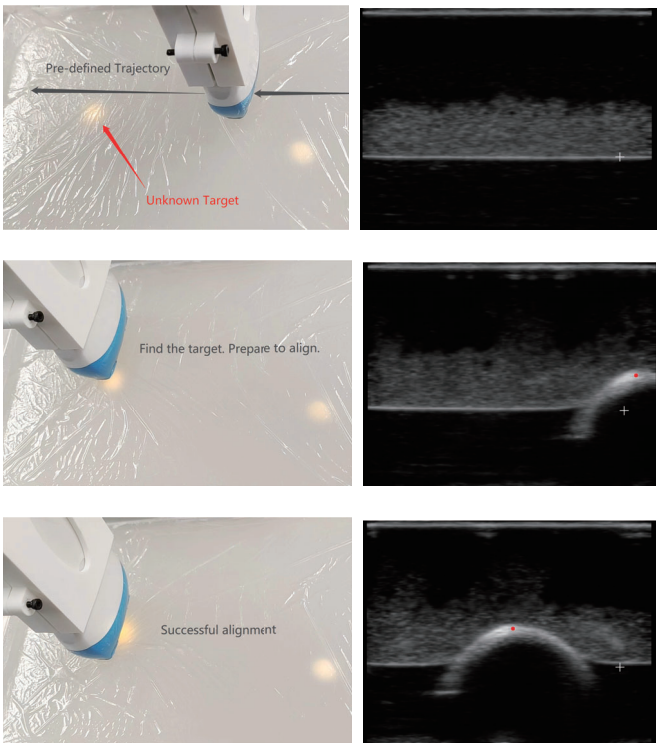

Fig. 9: The US scanner autonomously scans the plane following a predefined S-shaped trajectory until it encounters a sphere. The feature points and coordinates of the sphere are continuously outputted in the deep learning model, which subsequently operates the robotic arm for centering and positioning.

where n is the number of samples, y_i is the actual value for the i-th sample, $\hat{y}_i$ is the predicted value for the i-th sample, serving as a holistic gauge of tracking precision. Across three distinct trials, target forces of 15N, 20N, and 16N were implemented, based on the rationale that, we found that forces around 15N generally yield the best imaging results while also ensuring that subjects do not experience discomfort. The resulting RMSE errors for these trials were 4.23N, 2.54N, and 3.81N respectively fig. 8. This multi-dimensional approach was chosen to understand how various force-velocity pairings influence the performance of our method. By simulating a spectrum of real-world conditions, we aimed to delve deeper into the method's resilience and adaptability, ensuring its efficacy under different operational scenarios. Our findings unravel some compelling patterns. A direct correlation was noticed between the magnitude of the target force and the force tracking error: higher target forces often yielded smaller errors. On the contrary, elevated velocities were commonly associated with an uptick in the tracking error, signifying a more intricate relationship between force and velocity in tracking tasks. Specific force and velocity scenarios necessitate meticulous parameter tuning to harness optimal performance. Our experimental insights thus underscore the need for a robust method, adaptable across diverse operational scenarios, with a strong emphasis on fine-tuning its parameters. In the course of our research, we rigorously tested and validated our system's capabilities in two primary dimensions: tracking accuracy and force control. Through simulation environments, we went beyond mere validation, opting for comparative analyses that pitched our modulated controller against benchmark systems. These comparative evaluations underscored the tangible enhancements our controller brought to the table in terms of performance metrics.

Finally, we conducted a comprehensive scanning simulation test to ascertain whether our entire system could achieve the functionalities we anticipated. As illustrated in fig. 9, we allowed it to autonomously scan on a plane following a pre-defined trajectory until it encountered the sphere. The shape of the sphere was distinctly identified after processing through deep learning algorithms, and subsequently, it was centered and localized. The seamless operation throughout this process confirms the system's robust functionality.

When transitioning to real-world tests with the robotic arm, we delved deep into the intricacies of force dynamics. Not only did we analyze the forces at play, but we also implemented precise control mechanisms to modulate them as per our requirements. Coupling this with the observed trajectories, it became evident that our system is not just theoretically sound but practically adept. This culminated in our system being proficient in handling and completing fundamental scanning tasks.

IV. CONCLUSIONS

In our investigation, we harnessed two principal techniques for ultrasound-guided robot navigation. First, we incorporated the Residual Network (ResNet) architecture, recognized for its proficiency in discerning tissue anomalies and object localization. Thereafter, we embraced a time-invariant dynamical system (DS) approach to optimize robots in conducting ultrasound-guided scans. We delineated motion into

two phases: free motion and contact motion. The trajectory design for free motion underwent rigorous simulation tests, confirming our system's adept tracking prowess. Conversely, contact motion primarily emphasizes force regulation. Our empirical analysis juxtaposed the actual and anticipated force levels, revealing that our control mechanism aligns well with our objectives.

While our current research has made strides, it's evident that our reliance on planar scans presents certain limitations. These scans, though efficient for specific applications, fall short when faced with the complexities of scanning the human body or other irregular shapes. In our experiments, we utilized planar scans primarily to test foundational capabilities like target tracking, force control, and scan positioning. Recognizing these constraints, we're setting our sights on the future, aiming to augment our system with an enhanced visual module [19]. This enhancement is not just about handling more intricate scanning tasks; it's about elevating the precision of our scans to new heights. As we pivot towards this more sophisticated scanning approach, especially for applications like detailed human body scanning, we're also exploring the potential of refining our nominal dynamical system through demonstration learning techniques [20], [21]. This approach, we believe, might offer a more intuitive and adaptive edge over traditional methods [22]. Furthermore, as we embrace these advanced capabilities, our research methodology is poised for transformation. We're actively considering the integration of cutting-edge techniques like Support Vector Regression (SVR) and demonstration learning [23], equipping our system to tackle the challenges of non-planar and intricate scans with high accuracy.

REFERENCES

[1] A. M. Okamura, M. J. Matarić, and H. I. Christensen, "Medical and health-care robotics," *IEEE Robotics & Automation Magazine*, vol. 17, no. 3, pp. 26–37, 2010.

[2] A. Duan, M. Victorova, J. Zhao, Y. Sun, Y. Zheng, and D. Navarro-Alarcon, "Ultrasound-guided assistive robots for scoliosis assessment with optimization-based control and variable impedance," *IEEE Robotics and Automation Letters*, vol. 7, no. 3, pp. 8106–8113, 2022.

[3] K. Li, Y. Xu, and M. Q.-H. Meng, "An overview of systems and techniques for autonomous robotic ultrasound acquisitions," *IEEE Transactions on Medical Robotics and Bionics*, vol. 3, no. 2, pp. 510–524, 2021.

[4] A. Duan, C. Yang, J. Zhao, S. Huo, P. Zhou, W. Ma, Y. Zheng, and D. Navarro-Alarcon, "Safe learning by constraint-aware policy optimization for robotic ultrasound imaging," *IEEE Transactions on Automation Science and Engineering*, pp. 1–12, 2024.

[5] Q. Huang, J. Lan, and X. Li, "Robotic arm based automatic ultrasound scanning for three-dimensional imaging," *IEEE Transactions on Industrial Informatics*, vol. 15, no. 2, pp. 1173–1182, 2019.

[6] K. Li, Y. Xu, and M. Q.-H. Meng, "An overview of systems and techniques for autonomous robotic ultrasound acquisitions," *IEEE Transactions on Medical Robotics and Bionics*, vol. 3, no. 2, pp. 510–524, 2021.

[7] J. E. Colgate and N. Hogan, "Robust control of dynamically interacting systems," *International journal of Control*, vol. 48, no. 1, pp. 65–88, 1988.

[8] S.-L. J. Hu and H. Li, "Simultaneous mass, damping, and stiffness updating for dynamic systems," *AIAA journal*, vol. 45, no. 10, pp. 2529–2537, 2007.

[9] M. Victorova, M. K.-S. Lee, D. Navarro-Alarcon, and Y. Zheng, "Follow the curve: Robotic ultrasound navigation with learning-based localization of spinous processes for scoliosis assessment," *IEEE Access*, vol. 10, pp. 40216–40229, 2022.

[10] M. Tirindelli, M. Victorova, J. Esteban, S. T. Kim, D. Navarro-Alarcon, Y. P. Zheng, and N. Navab, "Force-ultrasound fusion: Bringing spine robotic-us to the next "level"," *IEEE Robotics and Automation Letters*, vol. 5, no. 4, pp. 5661–5668, 2020.

[11] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 770–778, 2016.

[12] S. Ioffe and C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," in *International conference on machine learning*, pp. 448–456, pmlr, 2015.

[13] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," *Advances in neural information processing systems*, vol. 25, 2012.

[14] W. Amanhoud, M. Khoramshahi, M. Bonnesoeur, and A. Billard, "Force adaptation in contact tasks with dynamical systems," in *2020 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 6841–6847, IEEE, 2020.

[15] K. Kronander and A. Billard, "Passive interaction control with dynamical systems," *IEEE Robotics and Automation Letters*, vol. 1, no. 1, pp. 106–113, 2015.

[16] S. S. M. Salehian and A. Billard, "A dynamical-system-based approach for controlling robotic manipulators during noncontact/contact transitions," *IEEE Robotics and Automation Letters*, vol. 3, no. 4, pp. 2738–2745, 2018.

[17] A. Duan, I. Batzianoulis, R. Camoriano, L. Rosasco, D. Pucci, and A. Billard, "A structured prediction approach for robot imitation learning," *The International Journal of Robotics Research*, vol. 43, no. 2, pp. 113–133, 2024.

[18] A. Duan, R. Camoriano, D. Ferigo, Y. Huang, D. Calandriello, L. Rosasco, and D. Pucci, "Learning to sequence multiple tasks with competing constraints," in *2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 2672–2678, 2019.

[19] L. Hu, A. Duan, M. Li, A. Cherubini, L. Li, and D. Navarro-Alarcon, "Paint with the sun: A thermal-vision guided robot to harness solar energy for heliography," *IEEE Sensors Journal*, vol. 22, no. 18, pp. 18130–18142, 2022.

[20] E. Gribovskaya, S. M. Khansari-Zadeh, and A. Billard, "Learning nonlinear multivariate dynamics of motion in robotic manipulators," *The International Journal of Robotics Research*, vol. 30, no. 1, pp. 80–117, 2011.

[21] A. Lemme, K. Neumann, R. F. Reinhart, and J. J. Steil, "Neural learning of vector fields for encoding stable dynamical systems," *Neurocomputing*, vol. 141, pp. 3–14, 2014.

[22] P. Zhou, P. Zheng, J. Qi, C. Li, A. Duan, M. Xu, V. Wu, and D. Navarro-Alarcon, "Neural reactive path planning with riemannian motion policies for robotic silicone sealing," *Robotics and Computer-Integrated Manufacturing*, vol. 81, p. 102518, 2023.

[23] P. Zhou, J. Qi, A. Duan, S. Huo, Z. Wu, and D. Navarro-Alarcon, "Imitating tool-based garment folding from a single visual observation using hand-object graph dynamics," *IEEE Transactions on Industrial Informatics*, vol. 20, no. 4, pp. 6245–6256, 2024.