IEEE Robotics and Automation Letters (RA-L) paper, presented at ICRA 2026, Vienna, Austria. Cite as RA-L paper.

A Soft-Rigid Hybrid Robot-Assisted Feeding System with A Tendon-Driven Continuum Robot

Jingyi Chen[1], Quecheng Qiu[1], and Jianmin Ji[1,2]

***Abstract*—Active delivery of food to a human mouth in a controlled and safe manner remains a key challenge for robot-assisted feeding systems (RAFSs). Existing RAFS designs struggle to simultaneously achieve efficiency and safety: rigid manipulators offer fast and accurate motion but risk hazardous contact, while soft robots provide passive compliance at the cost of limited speed or workspace. To meet the specific demands of feeding tasks, we design a tendon-driven continuum robot that allows precise orientation control of the utensil while exhibiting strong passive compliance in position. Integrating it with a 6-DoF rigid robot for fast and long-range positioning, we propose a hybrid RAFS architecture that achieves safe, efficient, and accurate food delivery. Controlling a passive-compliant RAFS to acquire various food is non-trivial: physical modeling struggles with complex interactions between soft robot and food, while typical imitation learning methods lead to discontinuous or distorted movements out of the passive deformation. To handle this, we design a pose-torque learning policy that enables the soft and rigid robots to generate coherent and synchronized movements, offering a case-specific solution to the long-standing challenge of soft robot imitation learning. Experiments show that our method achieve a food acquisition success rate of 76.7%, while user tests with 14 volunteers confirm user preference, marking our RAFS as a practical step toward safe and efficient robotic feeding.**

***Index Terms*—Soft Robot, Imitation Learning, Robot-Assisted Feeding System**

I. Introduction

ACTIVE Robot-Assisted Feeding System (RAFS) is a key application of personal assistant robots, aiming to autonomously acquire food and deliver it to the user's mouth [1], [2]. A RAFS must offer fast motion speed, high precision, sufficient workspace, and above all, safety during human-robot interaction [3]. Most existing RAFSs rely on motor-driven rigid robots which offer speed, accuracy, and reach, but their rigid construction and inherent mechanical stiffness introduce collision risks during human-robot interaction [4]. Although algorithmic safeguards like force threshold monitoring [5] or anomaly detector [6], [7] may enhance safety, these control methods face critical time-latency in sudden accident if user sneezes or coughs.

Manuscript received: May 10, 2025; Revised August 22, 2025; Accepted October 18, 2025. This paper was recommended for publication by Editor Cosimo Della Santina upon evaluation of the Associate Editor and Reviewers' comments. The work is partially supported by Hunan Province Major Scientific and Technological Project No. 2024QK2001, National Key R&D Program of China No. 2023YFB4704500, and National Natural Science Foundation of China No. 62332016.

[1] School of Computer Science and Technology, USTC, 230026, China.
[2] Institute of Artificial Intelligence, Hefei Comprehensive National Science Center, 230088, China (e-mail: Chenjingyi097@mail.ustc.edu.cn, qiuqc@mail.ustc.edu.cn, jianmin@ustc.edu.cn).

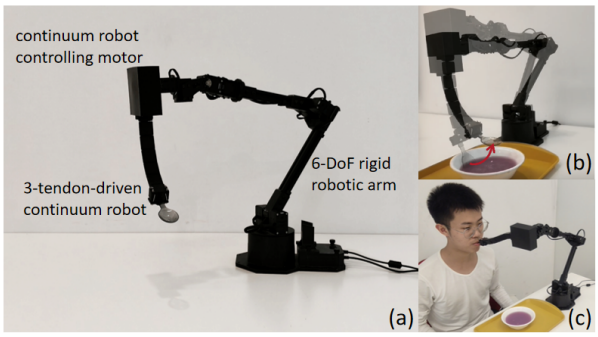


Fig. 1. The overview of our RAFS. (a) A three-tendon-driven continuum robot is actuated by a motor and connected with a 6-DoF rigid robotic arm. (b) Using coordinated control, the rigid and soft robot move synchronously to acquire food. (c) The RAFS actively and safely delivers food into the user's mouth without requiring user's movement.

On the other hand, soft robotics has emerged as a significant research area [8], with a growing exploration of feeding tasks. Soft robots exhibit passive compliance: they deform naturally under external forces due to the intrinsic material and structural design [4]. This allows them to react quickly and passively to unexpected contact without relying on sensing or active control, making them suitable for the safety demands of RAFS. However, the mechanical designs of soft robots also present challenges in feeding applications, including complex control mechanisms, slower movements, or limited workspaces. Recent efforts have explored the application of soft robots in feeding tasks [9], [10], showing improved safety for users. However, these designs suffer from insufficient speed or a limited workspace, rendering them impractical as effective RAFS solutions.

In this paper, our first goal is to develop a more capable RAFS architecture tailored to the demands of feeding tasks. We begin by analyzing the motion requirements of utensil-based food delivery and design a three-tendon-driven continuum robot. This soft robot enables precise orientation control of the utensil while maintaining strong compliance in position, allowing it to passively conform to sudden collision. By combining it with a 6-DoF rigid robot that provides fast, accurate motion and a large workspace, we propose a hybrid RAFS that effectively balances accuracy, speed, reach and safety, addressing the core requirements of RAFS.

However, this structural design introduces new challenges for successfully acquiring food. The interactions between the utensil and food can vary significantly, such as scooping, spearing, and dipping, each involving distinct contact dy-

IEEE Robotics and Automation Letters (RA-L) paper, presented at ICRA 2026, Vienna, Austria. Cite as RA-L paper.

namics. This intricate process entails numerous variations, making it impractical to explicitly model every possible food-interaction task. Furthermore, accurately estimating the state of a soft robot during interactions remains difficult, which complicates the application of classical control methods [11].

To bypass the complexity of modeling and task definition, recent advances in imitation learning(IL) propose to learn end-to-end policies directly from expert demonstrations. By leveraging closed-loop visual feedback, these methods can handle long-horizon manipulation tasks while maintaining task-level precision [12], [13]. Nevertheless, due to their unique structure and control methods, soft robots are less commonly used with imitation learning [14]. For our tendon-driven continuum robot, its compliance makes manual guidance teaching infeasible, while its elastic structure causes traditional IL methods to produce discontinuous movements.

Therefore, the second focus of our work is to design an imitation learning method suitable for our soft robot and apply it to the entire hybrid RAFS. To collect expert demonstrations, we leverage VisionPro [15] to track human hand and develop a customized motion mapping strategy that translates the motions into teleoperation commands for our hybrid RAFS. For policy learning, we build upon the state-of-the-art imitation learning method, ACT [12], which uses few data for effective policy learning. To overcome the challenge of discontinuous motion caused by stepwise position control of the soft robot, we adopt a torque-based control strategy that enables smooth and consistent behavior. Observing that torque can remain stable across multiple control frames, we design a hybrid imitation learning framework: the rigid robot is trained using conventional joint-position learning, while the soft robot is controlled in torque mode through a specialized algorithm. This decoupled yet synchronized design ensures coherent motion across both robots during food acquisition tasks. After filtering 50 optimal demonstrations as expert data, including the RAFS's motion and RGB-images from 2 cameras, the trained policy enables our robot to autonomously acquire food in consistent environments.

For validation, we design two food interaction scenarios and apply our imitation learning method to real world experiments. These tests compare the performance of our soft-rigid hybrid RAFS with a rigid-only RAFS, demonstrating the feasibility and effectiveness of our approach. The experiments highlight the advantages of hybrid RAFS in handling different tasks and ensuring safety during interactions. Furthermore, we conduct user studies with 14 volunteers divided into two groups to evaluate the systems, and the results reveal a strong preference for the soft-rigid hybrid RAFS over the traditional rigid design.

Our main contributions are summarized as follows:

- We propose a novel soft-rigid hybrid RAFS that combines a tendon-driven continuum robot with a 6-DoF rigid robot. The system enables fast, long-range motion through the rigid arm, while the soft robot offers precise orientation control with strong passive compliance in position, effectively acquiring food and absorbing sudden interacting collisions.
- We develop a pose-torque imitation learning framework tailored to the hybrid RAFS. Our method enables synchronized and continuous movements across the soft and rigid robots, offering a solution example for soft robot imitation learning.

II. Related Work

RAFSs involve a wide range of disciplines, including mechanical design, computer vision, speech recognition, human-robot interaction, feeding process design, etc [3]. In this paper, we mainly focus on the hardware design paradigms, safety-critical interaction, approach with soft robots, and learning-based control methods.

A. RAFS Hardware design paradigms

Existing RAFS hardware designs can be categorized based on the specifications of the robotic arms employed. Many studies utilize larger robotic arms, often integrated with wheelchair-like structures, prioritizing functional comprehensiveness [5], [16]–[21]. Existing work established face recognition [2] and food detection [22] as standard features. Advanced implementations incorporate speech command interfaces [23] or safety monitoring modules [7].

While larger robotic arms generally come with higher costs, a more cost-effective design approach involves using compact robotic arms situated on tabletops [1], [24]–[27]. These architectures demonstrate comparatively constrained performance in critical metrics such as operational workspace and payload capacity. Parikh et al. [25] systematically compare representative RAFS implementations, including their mechanical structures, functional features, technical limitations and estimated manufacturing costs.

B. Human-Robot Feeding Interaction

Mouth-related physical interaction constitutes the most critical phase in feeding robotics, where centimeter-level positioning errors or delayed response may lead to injury. Some systems established baseline force thresholds with emergency stop mechanisms [17]. Recent advancements achieve compliance adaptation through impedance control [16]. Particular researches focus on how human bites the spoon [28]–[30]. Nevertheless, conventional control strategies face inherent limitations: time-delayed response during sudden contacts may lead to potential risks, while sensor noise frequently trigger false alarms.

C. Feeding With Soft Robots

The inherent limitations of active compliance control may lead to potential risks. As a result, several research efforts have explored the use of soft robot for feeding tasks. While soft robotics may provide enhanced safety features, its control mechanisms are relatively complex due to the variety of unique movements, which make the control of soft robots particularly challenging [4]. Guan et al. [9] designed a trunk-like soft robot with extensive mobility, yet it has a slow motion speed, taking 90 seconds to complete a feeding action. Xie et al. [10] offers a more compact design, but it has a limited workspace and restricted end-effector DoF, making it difficult to execute the

IEEE Robotics and Automation Letters (RA-L) paper, presented at ICRA 2026, Vienna, Austria. Cite as RA-L paper.

entire feeding task independently. Keely et al. [31] developed a spoon consisting of a hoop and deformable sheet that prevents dental collisions, yet its hoop is made of rigid plastic.

D. Imitation Learning

Imitation learning enables efficient robotic skill transfer through human demonstrations. State-of-the-art methods address critical challenges: Action Chunking Transformers (ACT) [12] mitigate error propagation via temporal action segmentation, while diffusion policies [32] generate multi-modal actions through gradient-based denoising. Foundation models like RDT [33], [34] further enhance cross-domain generalization via large-scale pre-training.

Since our soft-rigid hybrid robot differs significantly from the common robotic arms with a gripper as the end effector, and since human eating data are difficult to collect in large quantities, we design a method of ACT style for training, which is more friendly to imitation learning with few data.

However, imitation learning remains underdeveloped in soft robotics compared to rigid robots, due challenges in demonstration flexibility, variable kinematics, and stochastic material behaviors [14], [35]. This necessitates the development of tailored imitation learning frameworks specifically for our soft-rigid hybrid RAFS architecture.

III. Approach

In this section, we first present the design of our soft robot, detailing its mechanical characteristics and how it is integrated with the rigid robotic arm. Subsequently, we elaborate on the motion design, demonstrating how our RAFS achieves human-like motion patterns for food acquisition and feeding operations. Finally, we introduce the imitation learning framework that acquires expert trajectories from human demonstrations, coupled with a specialized pose-torque imitation learning framework for RAFS manipulation.

A. Soft-Rigid Hybrid Mechanism Design

Our mechanism design begins with kinematic analysis of spoon-mouth interaction during feeding. Through observations detailed in III-B, we established two critical requirements: precise angular control of the spoon, and enough compliance in position, which motivated our design to be a 3-tendon-driven uniform-diameter continuum robot with only one motor. In the following three paragraphs, we answer three critical questions of our design: why do we select tendon-driven actuation, why is it uniform-diameter, and how to control the three tendons with one motor.

The actuation modality fundamentally determines soft robot performance [11]. While pneumatic or shape-memory alloy systems exist, we adopt tendon-driven actuation based on two critical considerations: faster response time enabling low-latency control during feeding, and lower hardware cost [36], which align with RAFS's need for responsiveness and affordability.

Through analytical modeling and calculation, we observe that the tendon-driven uniform-diameter continuum robot exhibits a favorable property: when subjected to moderate bending without twisting, the orientation can be approximated from tendon lengths, though positional determination remains infeasible [37], [38]. This characteristic aligns with our aforementioned requirements, thus motivating our design of a uniform continuum robot shown in Fig. 2(a). Building upon the joint model in Fig. 2(b), we develop an orientation estimation algorithm, with its computational error experimentally evaluated in Exp. IV-A. With a length of 21 cm, the soft robot shows a positional deviation of over 15 cm at its tip when the tendon lengths remain unchanged(head movement $<$15 cm during sneezing [39]).

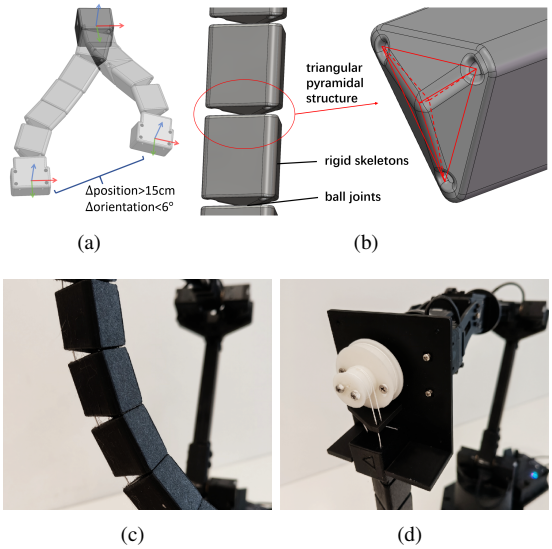


Fig. 2. (a) Large variations in position and approximately determined orientation under fixed tendon lengths. (b) Triangular pyramidal joint structure: rigid skeletons connected by ball joints. (c) Soft robot bending deformation. (d) Coaxial fixed-pulley (1:2 ratio) driven by a motor.

For such design, Webster et al. [40] suggest constant total tendon length during bending. But for our discretized structure model, mathematical analysis reveals that the minimum sum of tendon lengths reaches 94% of its maximum value, also validated in Exp. IV-A. Considering this quasi-constant characteristic, we develop the coaxial fixed-pulley configuration shown in Fig. 2(d): By winding two tendons around a fixed pulley (radius=1) and the third tendon inversely on the larger pulley (radius=2), single-motor actuation achieves unidirectional bending (note that the soft robot still has 3-DoF passive compliance).

We use Thermoplastic Polyurethane (TPU) for the skeleton and joints, along with strong, non-stretchable threads for the tendons, following [36].

B. Human-like Motion Design

The combination of our rigid and soft robot creates infinite solutions for achieving 6-DoF target spoon poses. Therefore, it is necessary to design a reasonable motion strategy. For the motion of delivering food into the mouth, we analyze through video observations and divide the feeding motion into three sequential steps (Fig. 3):

In the first step, the spoon approaches the mouth from a slightly lower position; secondly, it's inserted into the mouth;

IEEE Robotics and Automation Letters (RA-L) paper, presented at ICRA 2026, Vienna, Austria. Cite as RA-L paper.

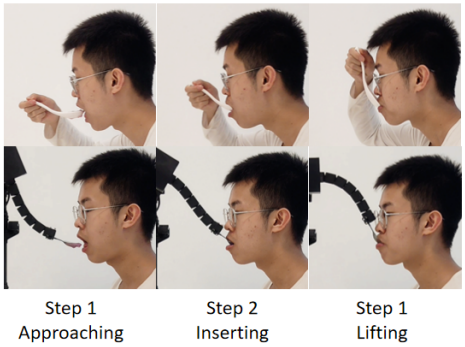


Fig. 3. Comparative side view of human-operated spoon feeding and our RAFS-assisted feeding operation. Trajectory similarity is visually evident, and the contact force during spoon-mouth interaction is also imitated by this motion design.

and finally, its lifted while maintaining upper-lip contact, then pulled out once it reaches an approximate angle of 60°. In this process, users may exhibit an error of 10°, while larger deviations may cause food spillage. To address safety and comfort during this human-mouth interaction, and to reduce the motion range and speed of the rigid robot, we propose the following robotic motion: the rigid robot lowers and rotates its end, while the soft robot performs synchronized bending. Below, we elucidate the advantages of this motion design.

In the first step, as the soft robot stretches, the spoon extends forward from a lower position into the mouth while simultaneously rising. In the second step, as the spoon fully enters the mouth, safety becomes paramount, so our soft robot extends straight, maximizing its range of passive compliance. During the final step, as the spoon is leaving the mouth, the soft robot bends significantly. The increased bending at this stage provides ample elastic force from the joints to assist the resistance of the spoon against the upper lip, ensuring a comfort feeling of contact between the upper lip and the lower lip.

For the motion of human manipulation of utensils during food interaction, we design it using similar analysis. When reaching for food with utensils, our RAFS maintains the soft robot in an extended configuration while coordinating rigid arm movements. During food scooping actions, the system achieves matching pitch angles through soft robot bending, while synchronously driving the rigid arm along corresponding paths. This controlled bending strategically reduces downward compliance, thus enhancing grasping stability during food transfer operations. This motion strategy for acquiring food is also validated in Exp. IV-C.

C. Unified Imitation Learning Framework

It's challenging to accurately estimate the state of a soft robot during interaction with diverse food. Though imitation learning might circumvent the difficulties, it faces new challenges: the difficulty of data collection, and the unique structural complexities of our soft robot.

For data collection, the deformation of soft robots under external force makes kinesthetic teaching difficult. Following the human-like motion design ideas just discussed above, we design a human-robot motion mapping strategy for teleoperation. We employ VisionPro [15] to capture and map real-time human hand movements to our RAFS, detailed in Alg. 1. $S.T$, $S.P$, and $R.P$ represent soft robot motor torque, spatial position of soft robot, and rigid robot pose, respectively. The utensil's pose is derived from the volunteers' hand data under the assumption of a consistent grip posture. Since the volunteers sometimes failed to successfully acquire food during teleoperation sessions, we manually select qualified expert datasets from the raw recordings by applying the following selection criteria: 1) Successful food acquisition; 2) Absence of bowl displacement, food spillage, or other undesired incidents during operation; 3) Action completion within 30 seconds.

The core challenge lies in the distinct behavior of our soft robot: when applying position control, the elastic forces of the soft robot interfere with the intended motor control, causing redundant oscillations. This makes it difficult to use typical imitation learning methods, since predicting motor positions step by step would cause discontinuous movements in the soft robot. However, we observe that using torque control, the soft robot moves continuously as the torque can maintain unchanged values across multiple frames. Given this, we choose to learn the motor's torque for the soft robot. Specifically, we control the motors of rigid robot in position mode and control the motor of soft robot in torque mode.

Based on this insight, we propose a hybrid pose-torque imitation learning approach. This method maintains joint-position learning for the rigid robot, while introducing a specialized torque learning algorithm for the soft robot's motor, as detailed in Alg. 2 and illustrated in Fig. 5. $A_{t-k,t}$ denotes the IL predicted action over the most recent k+1 frames, each including $S.T_E$, $S.P_E$ and $R.P_E$ representing the expected $S.T$, $S.P$ and $R.P$. $S.P_C$ means the current $S.P$, while $R.P_{OUT}$ and $S.T_{OUT}$ are the control commands output to the RAFS. The rationale of Alg. 2 is briefly shown in IV-C without a detailed explanation.

Our imitation learning framework is illustrated in Fig. 4. We implement our policy as a conditional variational autoencoder (CVAE) [41] that generates a short-horizon action sequence conditioned on the current observations. Visual inputs are processed with separate ResNet-18 backbones, which are flattened into sequences. To retain spatial information, a 2D learnable positional embedding is added to each sequence [12]. Then the visual feature, the RAFS pose and torque form the input to a DETR-style Transformer encoder [42]. Synthesized features feed the decoder to predict motion sequences, after which we employ Alg.2 as post-processing.

Algorithm 1: Expert Data Collecting

1: **Input**: $SpoonPose(x,y,z,roll,pitch,yaw)$
2: $S.T = S.T(-\frac{\pi}{2}, pitch)$
3: $S.P = S.P(x,y,z,-\frac{\pi}{2}, pitch)$
4: $R.P = S.P + RelativePose$
5: **Output**: $R.P,\ S.T,\ S.P$

IEEE Robotics and Automation Letters (RA-L) paper, presented at ICRA 2026, Vienna, Austria. Cite as RA-L paper.

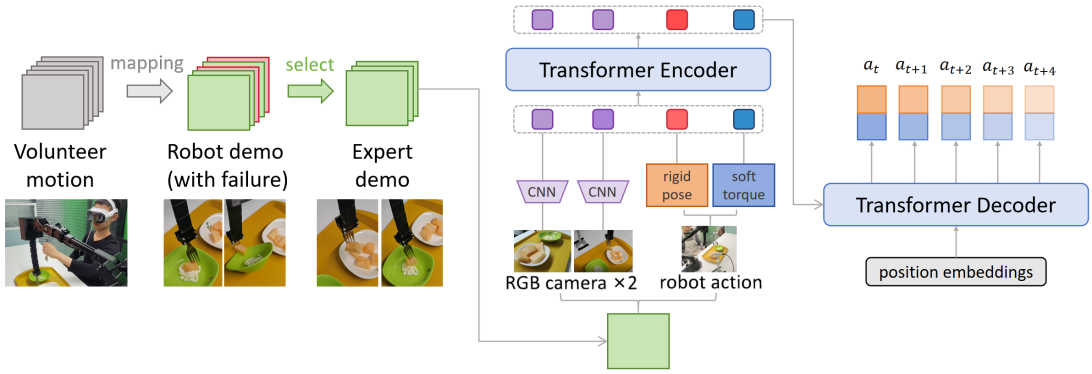


Fig. 4. Pipeline of our IL framework. VisionPro captures volunteer demonstrations, which are mapped to robot trajectories then selected for expert datasets. For each task, 4 demonstrators teleoperate the food-interaction phase until a total of 50 expert demonstrations are recorded. Multimodal data includes 2 camera RGB-images and RAFS motions. Images are encoded by CNN, and 2D positional embedding is applied to the feature sequence to keep spatial information. The RAFS motion is decomposed into rigid robot poses and soft robot torques. Synthesized features feed the decoder to generate motion sequences, and the post-processing is in Fig.5.

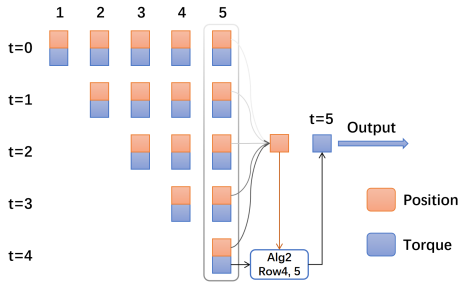


Fig. 5. The torque of soft robot controlling motor is adjusted based on the predicted position and torque.

Algorithm 2: Post-processing for IL Output

Input: $S.P_C$, $A_{t-k:t}$(*each including* $S.T_E$, $S.P_E$, $R.P_E$)
$R.P_{OUT} = \sum_{i=0}^{k} w_i A_{t-i}.R.P_E / \sum_{i=0}^{k} w_i$, *with* $w_i = e^{-i}$
$S.P_1 = \sum_{i=0}^{k} w_i A_{t-i}.S.P_E / \sum_{i=0}^{k} w_i$, *with* $w_i = e^{-i}$
$\Delta S.P = S.P_1 - S.P_C$
$S.T_{OUT} = A_t.S.T_E \times \left(1 + sgn(\Delta S_P) \times [\Delta S_P - 1]\right)$
Output: $R.P_{OUT}$, $S.T_{OUT}$

IV. Experiments

To systematically validate our RAFS, we conduct a hierarchical evaluation with six experiments. We begin with model verification in MuJoCo IV-A to confirm key mechanical assumptions III-A. Then, from a functional perspective, we assess the system's physical interaction characteristics through passive compliance IV-B, active deformation IV-C, and food acquisition tests IV-D. To validate the imitation learning method designed for our RAFS, we conduct comparative experiments against both a classical imitation learning method and a rigid-only RAFS setup IV-E. Finally, we conduct a user study with 14 volunteers, comparing our hybrid RAFS and a rigid-only RAFS in full feeding sessions to assess user experience and preference IV-F.

A. Model Validation through Simulation

To establish the theoretical foundation for subsequent experiments, we first validate two critical kinematic assumptions through physics-based simulation: 1) the approximate constancy of total tendon length ($\sum L = L_1 + L_2 + L_3 \approx constant$), and 2) the feasibility of orientation estimation through tendon length measurements. We develop the model of our soft robot in MuJoCo, and perturb it to reach various postures by applying random forces to the three tendons and the base of soft robot. Specifically, at each timestep(1s), an independent tensile force is sampled from a uniform distribution in 0-1N for each tendon, and torque is sampled from a uniform distribution in 0-0.5Nm for the base rotation (comparable to the working load in IV-C). During this perturbation process, we record the lengths of the three tendons and the sum of them, shown in Fig. 6. Meanwhile, we estimate the orientation based on the tendon length and compare it with the actual angular deviation, as shown in Fig. 6. E_α and R_α are the estimated and real pitch angles, while E_β and R_β are the estimated and real yaw angles. Roll angles are not involved since the soft robot cannot twist by tendon controlling.

The length summation analysis in Fig. 6 reveals remarkable consistency with $\sum L$ minimum value is 96.8 ± 0.4% of its maximum value across 30 sampled postures. This confirms our structural design effectively maintains near-constant total tendon length through, and the control method with one motor (Fig. 2(d)) is effective. The angular estimation error remains bounded within 5.7° across all test cases, with mean absolute error (MAE) of 0.7°. This confirms that the orientation can be estimated from tendon lengths.

B. Passive Compliance Test

To test and measure the passive compliance of our soft robot, we pause the soft robot at step 2 in feeding interactions, and test the relationship between its end's position offset and the external forces applied. We also show the effect during actual human-robot collision. The results and demonstration are depicted in Fig. 7.

IEEE Robotics and Automation Letters (RA-L) paper, presented at ICRA 2026, Vienna, Austria. Cite as RA-L paper.

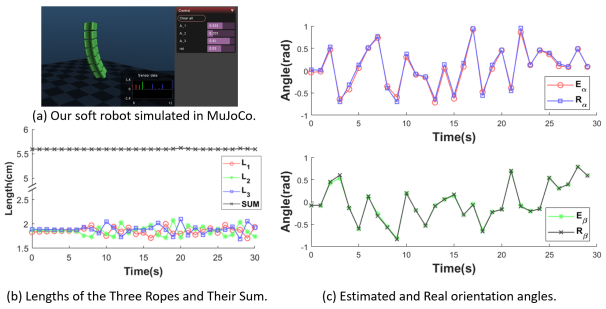


Fig. 6. Model validation by MuJoCo simulation across 30 perturbation trials. (a) We randomly perturbate the soft robot to check our kinematic assumptions. (b) Three-tendon length sum preserves 96.8% of maximum value; (c) Angular estimation demonstrates 0.7° MAE (5.7° maximum).

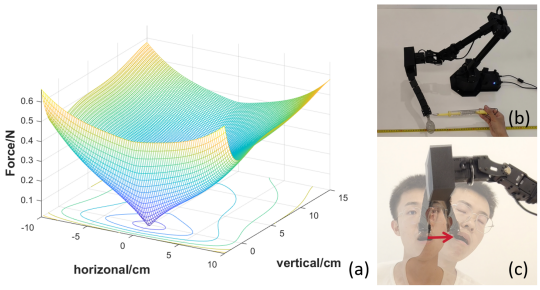


Fig. 7. Passive compliance test of our soft robot. (a) Relationship between the displacement of the utensil on the soft robot under external forces. (b) We conduct this test under quasi-static conditions. (c) The soft robot quickly yields to the collision when user coughs.

Fig. 7 shows that the soft robot exhibits remarkable compliance in various directions (e.g. 10 cm lateral deflection under 1 N external force), which significantly enhances the safety of the feeding process. It is crucial to emphasize that the passive compliance of our soft robot arises from its physical structure with almost no delay, rather than control methods. This is also why the compliance test graph of rigid robot is not shown: ignoring time-latency, rigid robot is capable of maintaining near-zero contact forces; but when it is dashed against, the contact force reaches a transient peak.

C. Active Deformation Control Verification

This experiment tests how our soft robot deforms under different payloads when actively controlling motor torque. It serves both as a payload working test and as a brief demonstration of the design rationale behind Alg. 2. We apply 4 different weights simulating food payloads to the end effector, while commanding tendon motor torques from 0.06 N·m to 0.2 N·m in 0.02 N·m increments.

As demonstrated in Fig. 8, a strong correlation emerges between commanded torque and motor position when payloads remain below 60 g (weight of an egg). For the remaining position error, we use Alg. 2 to perform torque adjustment, thereby providing more appropriate torque values based on the observed position discrepancies. Note that position here does not refer to the spoon's position, but rather to the position parameter returned by the motor controller, which in fact determines the spoon's orientation. This relationship between position and torque support for our torque-learning idea.

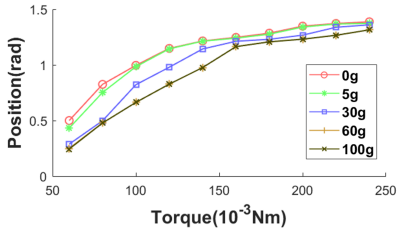


Fig. 8. Soft robot motor position under varying torque inputs and payloads. Under payload¡60g condition, commanded torque and motor position has strong correlations.

D. Food Acquisition Capability Test

To establish practical feeding competence, we evaluate our system's ability to acquire various foods through utensil-specific assessments, using teleoperation. Drawing inspiration from [16] and [22], we extend the test scope to eight typical food types in Fig. 9.

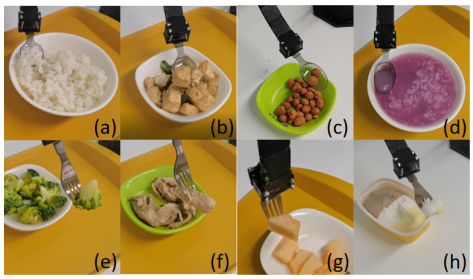


Fig. 9. The figure shows eight representative types of food that our robot is capable of handling. We use spoon for a) rice, b) tofu chunks, c) peanuts, d) porridge, and fork for e) broccoli, f) meat, g) fruit pieces, h) ice-cream.

This indicates that the robot can acquire common food items suitable for a spoon or fork, including small solids, liquids, viscoelastic soft foods, etc.

E. Imitation Learning Performance

To validate the effectiveness of our hybrid learning framework, we conduct comparative experiments on two representative feeding tasks shown in Fig. 10: 1) Scooping porridge from a bowl, 2) Spearing a piece of fruit and dipping into salad dressing. These two tasks involve complex scenarios, including utensil switching, three different food states, and various interactions. For this part of the experiment, we focus on food interaction, excluding active feeding. The experimental setup involves three configurations:

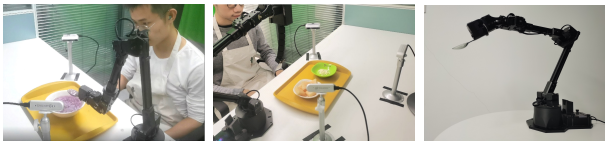
Group 1: Our soft-rigid hybrid RAFS, using the proposed method(III-C);

Group 2: Our soft-rigid hybrid RAFS, using typical position-based imitation learning, using same expert demonstration as Group 1;

Group 3: Only the 6-DoF rigid arm (WidowX-250S) with fixed utensils (Fig. 10(c)), using typical position-based imitation learning, using 50 dedicated demonstrations by directly mapping the motion of volunteers' utensil.

Performance is evaluated through 30 trials per task, using three assessments:

IEEE Robotics and Automation Letters (RA-L) paper, presented at ICRA 2026, Vienna, Austria. Cite as RA-L paper.



(a) (b) (c)

Fig. 10. Setup of the imitation learning experiments. (a)Task 1 requires simply scooping a spoon of porridge. (b)Task 2 involves spearing a piece of fruit and dipping it into salad dressing. (c) The rigid-only RAFS for comparison, using a WidowX-250s with torque control as the safety strategy.

- *Success Rate* (*SR*): The percentage ratio (%) of successfully delivering a spoonful of porridge or a piece of fruit dipped in salad dressing to the user within a specified time step during a single action.
- *Accident Rate* (*AR*): The percentage ratio (%) of severe accidents during food interactions, such as pushing the bowl too far or causing it to tip over.
- *Average Time* (*AT*): The average time (seconds) taken for a single action, representing the mean duration of successful actions recorded in the previous evaluation.

Note that the sum of SR and AR is not 100%, as sometimes the robot neither correctly acquires the food nor causes bad accident. Comparative evaluations between the soft-rigid hybrid system and conventional rigid robots reveal fundamental performance differences, as quantified in Table I. The hybrid RAFS with the pose-torque learning framework achieved the highest success rate and the lowest accident rate across both tasks. Notably, most accidents with the rigid-only RAFS were caused by knocking over the bowl, whereas the soft robot was able to avoid applying excessive force to the bowl when positional errors occurred.

In experiments, the RAFS moves slower than in demonstration, primarily due to the low operational frequency of our end-to-end control policy (e.g. 5 Hz), which causes slow and discrete movements. This trade-off prioritizes computational time for accurate visual perception and safe execution, and our future optimizations will focus on increasing policy frequency through model lightweighting and hardware upgrades to achieve smoother, human-like motion while maintaining safety.

TABLE I
FOOD INTERACTION EVALUATION RESULTS

	Porridge			Fruit		
Group	SR	AR	AT	SR	AR	AT
Group 1	**76.7**	**6.7**	17.4	**43.3**	**0**	24.5
Group 2	26.7	10	18.8	36.7	**0**	25.6
Group 3	50	26.7	**13.7**	23.3	13.3	**22.1**

F. Human-Robot Interaction Assessment

To establish preliminary validation of our RAFS's interactive performance under safety constraints, we conduct controlled experiments with 14 able-bodied volunteers randomly divided into 2 groups. One group used our hybrid RAFS, while the other tried the rigid-only RAFS with torque control as the safety strategy, shown in Fig. 10(c). Neither group of participants was exposed to the other system's mechanical structure. Other conditions were kept consistent: both groups completed 10 rounds of human–robot feeding interactions using peanuts as the food item, with realsense and Dlib toolkit for mouth detection. In both cases, the utensil followed the same trajectory as shown in Fig. 3, with the RAFS actively delivering food into the volunteer's mouth while participants remained still. Participants tried feeding interactions and evaluated each system using a 5-point Likert scale (1 = strongly disagree, 5 = strongly agree) across four typically used dimensions [3], [16], [25], [28], [43]:

- *Comfort* (*CMF*): Do you feel physically and psychologically comfortable during the feeding process? (1-5)
- *Adaptability* (*ADP*): Can you quickly synchronize with the robot's movements? (1-5)
- *Safety* (*SFT*): Do you feel confident that unintended collisions can be prevented during operation? (1-5)
- *Speed* (*SPD*): Does the motion pace match your natural interaction rhythm? (1-5)

Their rating results are shown in Table. II, and the comparison of the average scores for the two groups is displayed in Fig. 11.

TABLE II
THE SCORES GIVEN BY VOLUNTEERS

Robot	Volunteer	CMF	ADP	SFT	SPD
Soft	P1	5	4	5	5
	P2	3.5	3	4	4.5
	P3	4.5	4.5	5	3.5
	P4	5	5	4.5	4
	P5	5	4	5	4
	P6	4.5	5	4.5	4.5
	P7	4	5	4.5	4
Rigid	P8	4	3.5	3	3.5
	P9	3.5	3	3	4
	P10	3	3.5	3	4
	P11	4	3	4	5
	P12	3	2.5	3	3
	P13	3.5	3.5	3	4.5
	P14	3	3	3	4

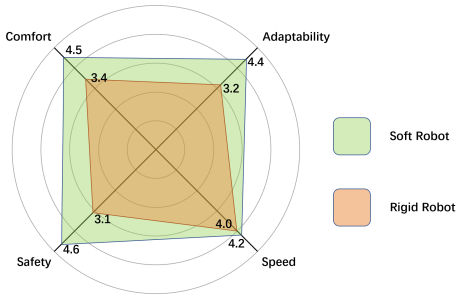


Fig. 11. The average scores given by volunteers across the four dimensions show clear preference of our hybrid RAFS.

The result demonstrates superior performance of our RAFS across all metrics, especially in terms of safety. It is plausible that, although participants in each group had no exposure to

IEEE Robotics and Automation Letters (RA-L) paper, presented at ICRA 2026, Vienna, Austria. Cite as RA-L paper.

the other group's RAFS structure, the appearance of the soft robot may have positively influenced their subjective ratings. Nonetheless, we believe that gaining users' trust based on appearance alone is also a valuable advantage.

V. Conclusion

We present a hybrid robot-assisted feeding system that combines a novel designed tendon-driven continuum robot with a 6-DoF rigid arm, effectively balancing safety and efficiency in feeding tasks. To address the challenges of control, we introduce a pose–torque imitation learning policy that generates coherent and synchronized motion across the rigid and soft robots. Experiments validate the system's practicality, highlighting its potential as a safe and efficient solution for robotic feeding. Additionally, intriguing challenges remain for future work. A key direction is determining how to leverage large human datasets for training structurally unique robots. Another is developing strategies that can dynamically adjust to the user's real-time state.

References

[1] W.-K. Song *et al.*, "Novel assistive robot for self-feeding," *Robotic Systems-Applications, Control and Programming*, pp. 43–60, 2012.

[2] A. Candeias *et al.*, "Vision augmented robot feeding," in *Proceedings of the European Conference on Computer Vision (ECCV) Workshops*, 2018, pp. 0–0.

[3] A. Nanavati *et al.*, "Design principles for robot-assisted feeding in social contexts," in *Proceedings of the 2023 ACM/IEEE International Conference on Human-Robot Interaction*, 2023, pp. 24–33.

[4] D. Rus *et al.*, "Design, fabrication and control of soft robots," *Nature*, vol. 521, no. 7553, pp. 467–475, 2015.

[5] A. Nanavati, "Achieving deployable autonomy through customizability and human-in-the-loop: a case study in robot-assisted feeding," in *Companion of the 2024 ACM/IEEE International Conference on Human-Robot Interaction*, 2024, pp. 136–138.

[6] D. Park *et al.*, "A multimodal execution monitor with anomaly classification for robot-assisted feeding," in *2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2017, pp. 5406–5413.

[7] ——, "A multimodal anomaly detector for robot-assisted feeding using an lstm-based variational autoencoder," *IEEE Robotics and Automation Letters*, vol. 3, no. 3, pp. 1544–1551, 2018.

[8] S. Kim *et al.*, "Soft robotics: a bioinspired evolution in robotics," *Trends in biotechnology*, vol. 31, no. 5, pp. 287–294, 2013.

[9] Q. Guan *et al.*, "Trimmed helicoids: an architectured soft structure yielding soft robots with high precision, large workspace, and compliant interactions," *npj Robotics*, vol. 1, no. 1, p. 4, 2023.

[10] Z. Xie *et al.*, "Soft robotic arm with extensible stiffening layer," *IEEE Robotics and Automation Letters*, vol. 8, no. 6, pp. 3597–3604, 2023.

[11] O. Yasa *et al.*, "An overview of soft robotics," *Annual Review of Control, Robotics, and Autonomous Systems*, vol. 6, no. 1, pp. 1–29, 2023.

[12] T. Z. Zhao *et al.*, "Learning fine-grained bimanual manipulation with low-cost hardware," *arXiv preprint arXiv:2304.13705*, 2023.

[13] C. Chi *et al.*, "Universal manipulation interface: In-the-wild robot teaching without in-the-wild robots," *arXiv preprint arXiv:2402.10329*, 2024.

[14] M. S. Nazeer *et al.*, "Imitation and reinforcement learning to control soft robots: a perspective," in *IOP Conference Series: Materials Science and Engineering*, vol. 1292, no. 1. IOP Publishing, 2023, p. 012010.

[15] Y. Park *et al.*, "Using apple vision pro to train and control robots," 2024. [Online]. Available: https://github.com/Improbable-AI/VisionProTeleop

[16] D. Park *et al.*, "Active robot-assisted feeding with a general-purpose mobile manipulator: Design, evaluation, and lessons learned," *Robotics and Autonomous Systems*, vol. 124, p. 103344, 2020.

[17] E. K. Gordon *et al.*, "An adaptable, safe, and portable robot-assisted feeding system," in *Companion of the 2024 ACM/IEEE International Conference on Human-Robot Interaction*, 2024, pp. 74–76.

[18] S. Leone *et al.*, "Design of a wheelchair-mounted robotic arm for feeding assistance of upper-limb impaired patients," *Robotics*, vol. 13, no. 3, p. 38, 2024.

[19] T. Bhattacharjee *et al.*, "Is more autonomy always better? exploring preferences of users with mobility impairments in robot-assisted feeding," in *Proceedings of the 2020 ACM/IEEE international conference on human-robot interaction*, 2020, pp. 181–190.

[20] A. Nanavati *et al.*, "Lessons learned from designing and evaluating a robot-assisted feeding system for out-of-lab use," in *Proceedings of the 2025 ACM/IEEE International Conference on Human-Robot Interaction*, 2025, pp. 696–707.

[21] R. Feng *et al.*, "Robot-assisted feeding: Generalizing skewering strategies across food items on a plate," in *The International Symposium of Robotics Research*. Springer, 2019, pp. 427–442.

[22] R. Liu *et al.*, "Adaptive visual imitation learning for robotic assisted feeding across varied bowl configurations and food types," *arXiv preprint arXiv:2403.12891*, 2024.

[23] A. Padmanabha *et al.*, "Voicepilot: Harnessing llms as speech interfaces for physically assistive robots," in *Proceedings of the 37th Annual ACM Symposium on User Interface Software and Technology*, 2024, pp. 1–18.

[24] E. K. Gordon *et al.*, "Adaptive robot-assisted feeding: An online learning framework for acquiring previously unseen food items," in *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2020, pp. 9659–9666.

[25] P. Parikh *et al.*, "Design and development of a low-cost vision-based 6 dof assistive feeding robot for the aged and specially-abled people," *IETE Journal of Research*, vol. 70, no. 2, pp. 1716–1744, 2024.

[26] M. Guo *et al.*, "Development a feeding assistive robot for eating assist," in *2017 2nd Asia-Pacific Conference on Intelligent Robot Systems (ACIRS)*. IEEE, 2017, pp. 299–304.

[27] F. Liu *et al.*, "I-feed: A robotic platform of an assistive feeding robot for the disabled elderly population," *Technology and Health Care*, vol. 28, no. 4, pp. 425–429, 2020.

[28] R. K. Jenamani *et al.*, "Feel the bite: Robot-assisted inside-mouth bite transfer using robust mouth perception and physical interaction-aware control," in *Proceedings of the 2024 ACM/IEEE International Conference on Human-Robot Interaction*, 2024, pp. 313–322.

[29] J. Ondras *et al.*, "Human-robot commensality: Bite timing prediction for robot-assisted feeding in groups," in *6th Annual Conference on Robot Learning*, 2022.

[30] T. Bhattacharjee *et al.*, "Towards robotic feeding: Role of haptics in fork-based food manipulation," *IEEE Robotics and Automation Letters*, vol. 4, no. 2, pp. 1485–1492, 2019.

[31] M. N. Keely *et al.*, "Kiri-spoon: A soft shape-changing utensil for robot-assisted feeding," *arXiv preprint arXiv:2403.05784*, 2024.

[32] C. Chi *et al.*, "Diffusion policy: Visuomotor policy learning via action diffusion," *The International Journal of Robotics Research*, p. 02783649241273668, 2023.

[33] S. Liu *et al.*, "Rdt-1b: a diffusion foundation model for bimanual manipulation," *arXiv preprint arXiv:2410.07864*, 2024.

[34] O. M. Team *et al.*, "Octo: An open-source generalist robot policy," *arXiv preprint arXiv:2405.12213*, 2024.

[35] K. Chin *et al.*, "Machine learning for soft robotic sensing and control," *Advanced Intelligent Systems*, vol. 2, no. 6, p. 1900171, 2020.

[36] Z. Wang *et al.*, "Spirobs: Logarithmic spiral-shaped robots for versatile grasping across scales," *Device*, 2024.

[37] M.Russo *et al.*, "Continuum robots: An overview," *Advanced Intelligent Systems*, vol. 5, no. 5, p. 2200367, 2023.

[38] P. Rao *et al.*, "How to model tendon-driven continuum robots and benchmark modelling performance," *Frontiers in Robotics and AI*, vol. 7, p. 630245, 2021.

[39] P. Bahl *et al.*, "Flow dynamics of droplets expelled during sneezing," *Physics of Fluids*, vol. 33, no. 11, 2021.

[40] R. J. Webster *et al.*, "Design and kinematic modeling of constant curvature continuum robots: A review," *The International Journal of Robotics Research*, vol. 29, no. 13, pp. 1661–1683, 2010.

[41] K. Sohn *et al.*, "Learning structured output representation using deep conditional generative models," *Advances in neural information processing systems*, vol. 28, 2015.

[42] N. Carion *et al.*, "End-to-end object detection with transformers," in *European conference on computer vision*. Springer, 2020, pp. 213–229.

[43] T. Bhattacharjee *et al.*, "A community-centered design framework for robot-assisted feeding systems," in *Proceedings of the 21st International ACM SIGACCESS Conference on Computers and Accessibility*, 2019, pp. 482–494.