2026 IEEE International Conference on Robotics and Automation (ICRA 2026)
June 1-5, 2026. Vienna, Austria

MINT: A Vision-based Soft Sensor for Mutual Integration of Normal Interaction Force and Texture Perception

Mohammad Rafiee Javazm[1], Siddhartha Kapuria[1], Ozdemir Can Kara[1], Sonika Kiehler[1], Rami Hamada[2], and Farshid Alambeigi[1]

***Abstract*— Inspired by the design of Vision-based Tactile Sensors (VTSs) and soft resistive strain sensors, in this paper, we introduce MINT: a vision-based soft sensor for Mutual Integration of Normal interaction force and Texture perception. MINT is a hybrid vision-based tactile sensor that simultaneously integrates normal force measurement with high-resolution texture perception. This unique sensor utilizes a soft resistive strain sensor between the Gel Layer and Mirror Layer of a typical VTS. By combining electrical and visual sensing modalities, MINT overcomes the limitations of existing resistive sensors and VTSs, offering a robust, efficient, and scalable solution for *direct* measurement of force and texture capture. To evaluate MINT's functionality, we first propose a unique design and fabrication procedure. Next, we conduct a series of experiments, evaluating its force and texture sensing capabilities through interactions with various rigid objects.**

I. INTRODUCTION

In recent years, vision-based tactile sensors (VTS) have emerged as a powerful tool for capturing rich tactile information in robotic and haptic applications [1]. These sensors leverage compact form factor and lower cost optical imaging techniques to analyze deformations at the contact interface in pixel level, providing detailed insights into surface texture, contact forces, or both. Based on the VTSs' primary function, they can be classified into three main categories:

(i) *Texture-capturing VTSs*, such as BRDF [2] or HySenSe [3]–[5], focus on capturing high-resolution texture imaging of surface, making them ideal for material classification and surface quality assessment. These sensors utilize transparent elastomers and structured lighting to enhance texture contrast without directly or indirectly measuring contact forces.

(ii) *Force-estimation VTSs* rely on embedded markers displacement or material deformation captured by a camera to *directly* infer normal and shear interaction forces. Examples include TacTip [6] or the introduced VTS in [7], which are widely used in robotic gripping and manipulation tasks.

(iii) *Hybrid VTSs*, such as GelSight [8], GelSlim [9], DTact [10], [11], and DIGIT [12], combine both visual texture perception and *indirectly* force sensing within a single vision-based tactile sensing framework using Artificial Intelligence (AI) and machine learning (ML) algorithms, providing a comprehensive understanding of contact interactions. While the fusion of texture and force sensing provides a more comprehensive tactile sensing solution, it also introduces several challenges. For example, training such systems requires collecting large-scale datasets, a process that is both time-consuming and labor-intensive. Furthermore, variations in silicone properties and fabrication inconsistencies necessitate individualized calibration and training for each sensor making its use cases very limited. On the other hand, several studies have demonstrated high-performance contact force detection using conventional sensing principles, including piezoelectric [13], resistive [14], capacitive [15], triboelectric [16], and optical [17] approaches. These soft proprioceptive sensors are widely utilized in robotic skin [18] and biomimetic sensors [19], further advancing robotic perception and interaction capabilities. Among these, due to their simplicity and cost-effectiveness, soft *resistive* sensors are widely utilized for *directly* measuring interaction forces and pressure. These sensors operate based on changes in their length and cross-sectional area, which directly influence their electrical resistance. For instance, wearable soft resistive sensors (e.g., [20]–[24]) use a liquid metal (eutectic Indium Gallium alloy (EGaIn)) injected into spiral-shaped microchannels embedded within a stretchable silicone substrate. When subjected to deformation due to external interactions, the sensor exhibits a change in electrical resistance, which results from the geometric alterations of the microchannels (i.e., cross section area).

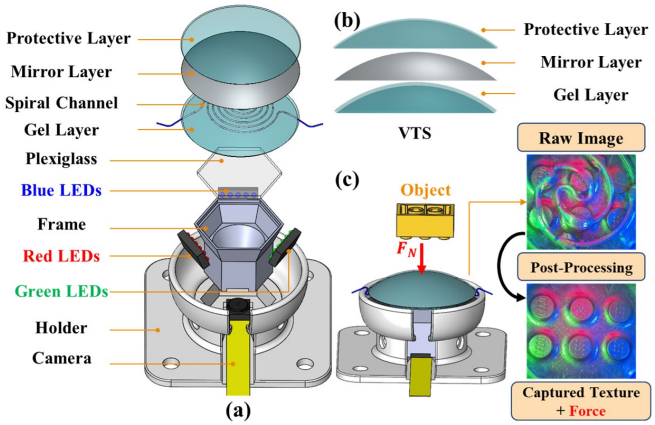

Fig. 1. Exploded view of the mechanical structure of (a) the proposed MINT sensor compared to (b) a traditional vision-based tactile sensor. (c) The MINT framework for simultaneous texture capture and direct measurement of interaction force.

Research reported in this publication was supported by the National Cancer Institute of the National Institutes of Health under Award Number R21CA280747.

1. M. R. Javazm, S. Kapuria, O. C. Kara, S. Kiehler, and F. Alambeigi are with the Walker Department of Mechanical Engineering and Texas Robotics, University of Texas at Austin, TX, USA.

2. R. Hamada is with the Chandra Family Department of Electrical and Computer Engineering and Texas Robotics, University of Texas at Austin, TX, USA. Email: mr62958@my.utexas.edu, {skapuria,ozdemirckara,sonikakiehler}@utexas.edu, rami.hamada@austin.utexas.edu and, farshid.alambeigi@austin.utexas.edu

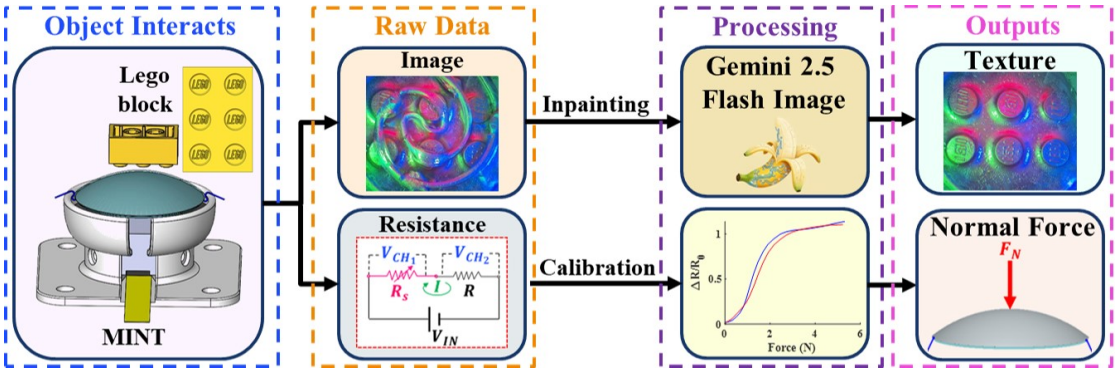

Fig. 2. Proposed framework for capturing texture and directly measuring normal interaction force. Top row: raw images captured by the camera are inpainted using prompt-driven generative inpainting of Gemini 2.5 Flash model to obtain textures comparable to those from conventional VTSs. Bottom row: measured resistance is converted to normal interaction force through a straightforward calibration process.

Despite their numerous advantages, soft resistive sensors, like other mentioned conventional tactile sensors, lack the ability to capture surface texture. This limitation restricts their applicability in tasks requiring both force estimation and texture recognition.

Drawing inspiration from VTSs and spiral-channel resistive tactile sensors, and aiming at addressing the limitation of either of these sensors, in this paper, we introduce MINT: a vision-based soft sensor for Mutual Integration of Normal interaction force and Texture perception. MINT is a hybrid vision-based tactile sensor that simultaneously integrates normal force measurement with high-resolution texture perception. As shown in Fig.1, it utilizes a soft strain sensor with a spiral channel between the Gel Layer and Mirror Layer of a VTS. Through this hybrid design, and with the framework shown in Fig.2, we are able to directly and simultaneously measure both texture and interaction force. As shown in Fig. 2, the interaction force is obtained directly through a simple calibration process by measuring resistance, while the object's texture is captured through an inpainting approach [25]–[28] applied to the raw texture images acquired by the camera. By combining electrical and visual sensing modalities, MINT framework overcomes the limitations of existing resistive sensors and VTSs, offering a robust, efficient, and scalable solution for tactile sensing. To evaluate MINT's functionality, we first propose a unique design and fabrication procedure. Next, we conduct a series of experiments, evaluating its force and texture sensing capabilities through interactions with various objects.

II. Design and Fabrication

As shown in Fig.1c, a high-resolution VTS for surface topography, such as BRDF [2], or HySenSe [3], typically consists of the following key components: (i) *Gel Layer*: Made from materials such as silicone or gel, this layer distributes applied forces across its surface while maintaining optical transparency to avoid obstructing the camera's field of view. (ii) *Mirror Layer*: A thin reflective or mirror coating, such as mirror paint or aluminum powder, is applied to prevent light leakage and enhance contrast for the vision system. (iii) *Protective Layer*: A soft material is placed over the reflective layer to protect it from wear and mechanical damage, ensuring durability and consistent performance. (iv) *Camera, LEDs, and Rigid Housing*: A rigid housing consisting of a transparent acrylic sheet to support the sensor, a frame to keep the sensor at the optimal focal distance of the camera, an array of Red, Green and Blue LEDs for controlled illumination, and a camera positioned beneath the acrylic sheet, facing the gel layer to capture its deformation in interaction with objects.

On the other hand, soft resistive sensors (e.g., [20]–[22]) use a liquid metal (e.g., eutectic Indium Gallium alloy (EGaIn)) injected into spiral-shaped microchannels embedded within a stretchable silicone substrate to measure the interaction force or pressure with an object. Inspired by the design of texture mapping VTSs and soft resistive sensors with spiral-shaped channel, to *directly* measure normal interaction forces while *simultaneously* capturing texture of an interacting object, we introduce MINT. As shown in Fig.1, MINT utilizes a soft strain sensor with Spiral channel [24] between the Gel Layer and Mirror Layer of a VTS. Through this hybrid design, we are able to directly and simultaneously measure both texture and interaction force.

To achieve the mentioned feature, MINT's design must satisfy the following key requirements: **(i)** *Simultaneous Sensing*: The sensor must concurrently capture high-resolution textural images and *directly* measure normal interaction forces by detecting changes in the cross-sectional area of the embedded spiral channel within the gel substrate; **(ii)** *Minimal Interference*: The integration should ensure high sensitivity in both texture perception and normal force measurement while minimizing mutual interference or performance degradation; **(iii)** *Minimized Obstruction of Field of View*: The spiral channel should be positioned in a way to minimize obstructing the camera's field of view. These design considerations ensure a soft and safe interaction, enabling precise texture acquisition while directly measuring normal interaction forces.

A. Design and Fabrication of MINT

Inspired by the spiral strain sensor [24] and the HySenSe sensor [3] as a VTS, we propose a hybrid design and fabrication approach. Integrating a soft spiral strain sensor with a traditional VTS design can be achieved through the following two different approaches:

(i) *Stacked Configuration*: In this approach, the VTS

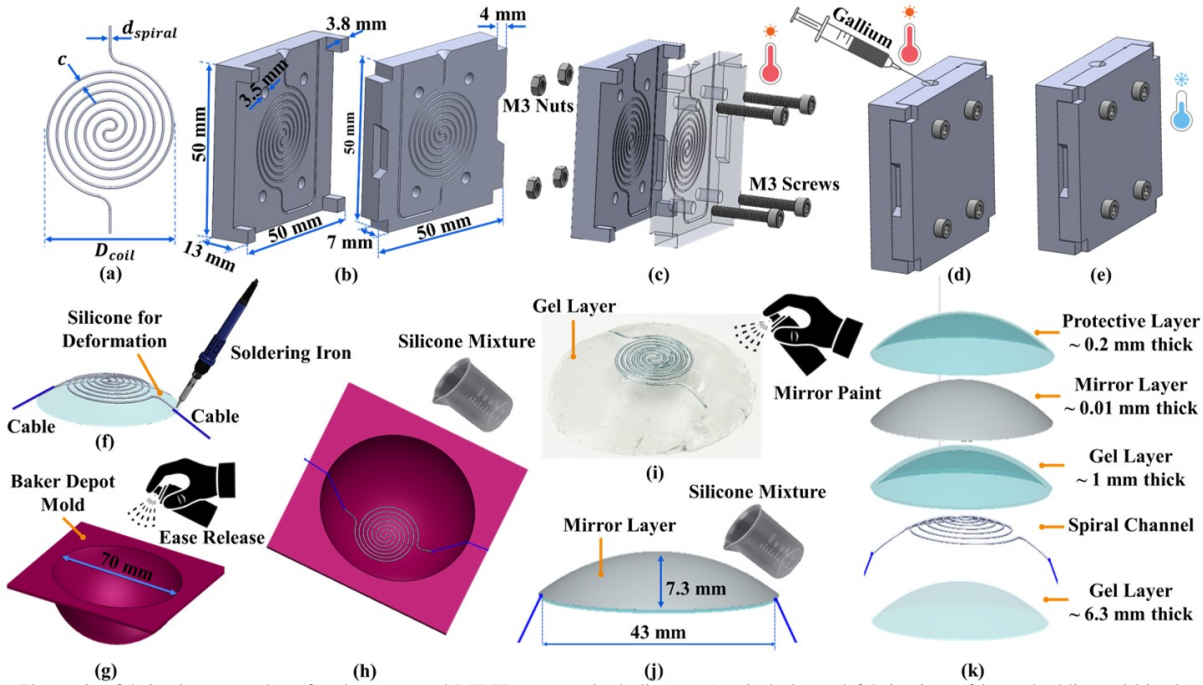

Fig. 3. The entire fabrication procedure for the proposed MINT sensor , including (a-e) spiral channel fabrication, (f-h) embedding within the substrate, (i-j) mirror-layer spraying, and (k) an exploded view of the sensor components.

sensor is fabricated first, and the spiral strain sensor is then stacked on top of it. This design does not obstruct the camera's field of view, as the strain sensor is positioned outside the imaging area, and it provides high sensitivity to normal interaction forces. However, the added layer increases the distance between the reflective layer and the sensor's surface, which can hinder texture capture.

(ii) *Hybrid Fabrication*: Alternatively, as shown in Fig. 1b, the spiral channel can be embedded directly within the gel layer of the VTS. Although this approach partially occludes the camera's field of view, careful optimization of the spiral channel's parameters can minimize obstruction while still maintaining high-quality texture capture. An inpainting method [25]–[28] also can then be applied to remove the spiral channel from the field of view and reconstruct the obstructed texture using the surrounding regions. In addition, the spiral channel can serve as a marker for further analysis, where its deformation during interaction with objects enables indirect force measurement through an AI-based model. By carefully balancing these trade-offs, the hybrid fabrication approach allows for simultaneous texture acquisition and direct normal force measurement, enhancing the overall functionality of the MINT framework. In the following section, we focus solely on evaluating the functionality of the hybrid fabrication method, as the stacked approach was found to reduce the effectiveness of texture mapping.

1) ***Design and Fabrication of MINT Holder****:* As shown in Fig.1, MINT features a dome-shaped deformable silicone layer that directly interacts with objects, along with a camera positioned to capture the deformation of the gel layer. The camera (Arducam 1/4 inch 5 MP sensor) is fixed to the rigid frame of the sensor 3D printed with PLA material (Raise 3D E2), which also supports a transparent acrylic layer. An array of Red, Green, and Blue LEDs (L-SMTD Mono-Color 150141RS63130, 150141GS63130, 150141BS63130) provides illumination, aiding in the recreation of 3D textural features when an object makes contact with the sensor.

2) ***Design and Fabrication of Spiral Channel****:* To fabricate the spiral-shaped gallium microchannels, as shown in Fig.3a, we designed and 3D-printed the cap and closure molds (Fig.3b) using high-resolution additive manufacturing approach. In particular, the molds were printed with a 3D printer (Anycubic M5s LCD printer) using high-clear resin (Anycubic). The mold design includes a guide for precise alignment of the cap and closure, M4 screw holes for secure clamping to prevent leakage during injection, and a vent line to facilitate air release and prevent air entrapment inside the mold. To assemble the mold, we aligned the cap and closure, then clamped them together by tightening M4 screws and nuts (Fig. 3c). The assembled mold was then placed inside an incubator and heated to 40°C to prevent premature solidification of the liquid gallium during injection. Simultaneously, the gallium was heated to 40°C, exceeding its melting point of 30°C, to ensure it remained in a fully liquid state. The liquid gallium was then injected into the mold (Fig. 3d). The transparency of the high-clear resin used for mold fabrication allowed real-time monitoring of the injection process, ensuring complete filling of the microchannels. Once injected, the mold containing the gallium was placed in a refrigerator to expedite the solidification process (Fig.3e). After solidification, the mold was carefully dismantled to extract the formed spiral-shaped gallium microchannel (see Fig.4).

To evaluate the performance of the MINT Framework under different conditions, three distinct spiral-shaped gallium microchannels with circular cross-sections were fabricated. These variations were designed to assess the influence of key parameters such as cross-sectional diameter (d_{spiral})

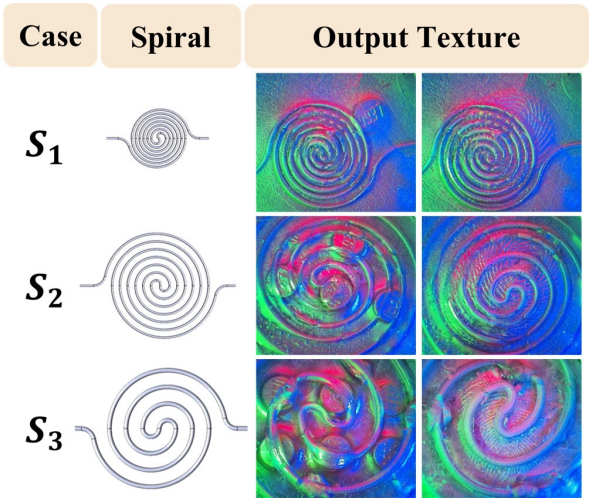

Fig. 4. Comparison of texture quality among three distinct MINT sensor configurations in detecting the Lego block pattern and fingerprint texture.

and coil gap distance (c) (see Fig. 3a). The specifications of these three spiral designs are summarized in Table I. Of note, these dimensions can be readily be changed for various applications. Since our sensor has a dome shape while the spiral channel is initially flat, we placed the flat spiral channel on top of a dome-shaped gel to conform it to the desired curvature (see Fig. 3f). The dome-shaped gel was created by pouring a silicone (P-565, Silicones Inc.) into a silicone cupcake mold (Baker Depot mold for chocolate) with a diameter of 70 mm and curing it. Once the spiral channel was properly adjusted to the dome shape, conducting wires were soldered to both ends of the spiral channel to enable resistance measurement of the microchannel. Figure 3 illustrates these steps.

*3) **Fabrication Procedure of the Gel Layer**:* To fabricate the spherical dome-shaped gel layer, we used a silicone mold (Baker Depot mold for chocolate) with a diameter of 70 mm. To embed the spiral channel within the spherical dome-shaped gel substrate, we first sprayed the mold's surface with Ease Release 200 (Mann Technologies) to prevent adhesion and facilitate easy removal of the cured silicone (Fig. 3g). Next, we placed the pre-shaped spiral channel with attached cables at the bottom of the mold (Fig. 3h). Of note, to ensure that the spherical shape of the mold remained unchanged during the curing process, we used Solo Crystal Clear Plastic Cups to stabilize and maintain its structure. Since the gel layer must be soft and transparent to avoid obstructing the camera's field of view, we fabricated the sensor using platinum-catalyzed transparent silicone rubber (P-565, Silicones Inc.). In this study, we used a 31:22:14 (A:B:C) mass ratio, where Part A is silicone, Part B is the activator, and Part C is phenyl trimethicone-softener (LC1550, Lotioncrafter) to achieve a softer structure. After thoroughly mixing these components, we degassed the mixture in a vacuum chamber to remove air bubbles and ensure optical transparency. The degassed mixture was then poured into the mold, covering the entire spiral channel (Fig. 3h). After curing for approximately 7 hours at room temperature, we removed the cured structure from the mold, as illustrated in Fig. 3i.

TABLE I
GEOMETRIC PARAMETERS OF THREE DIFFERENT SPIRAL CHANNELS

Sample	d_{spiral} (mm)	c (mm)	D_{coil} (mm)	**Texture** Quality	**Force** Quality
S_1	0.5	0.5	15.3	✗	✗
S_2	0.5	1.0	27.3	✓	✗
S_3	1.0	3.0	31.9	✓	✓

To prevent light leakage and enhance the reconstruction of 3D textural features, we sprayed the gel layer—now embedded with the spiral channel—with silver Glass Mirror Effect spray paint (Rust-Oleum Inc.) (Fig. 3i). The spraying process was repeated 10 times at 1-minute intervals to ensure complete coverage and uniform drying. Once the mirror layer was fully applied, we prepared another batch of silicone rubber gel using the same ratio and poured it over the mirror layer for its protection (Fig. 3j). As illustrated in Fig. 3k, the final structure of MINT sensor consists of four distinct layers:
(i) **Gel Layer**: Functions like a lens to magnify the texture.
(ii) **Spiral Channel**: Directly measures normal interaction forces by measuring resistance changes due to microchannel deformation.
(iii) **Mirror Layer**: Prevents light leakage and enhances the recreation of 3D textural features of objects.
(iv) **Protective Layer**: Shields the mirror coating from damage and degradation.

B. Occlusion Removal via Prompt-Based Inpainting

A key challenge in the usability of MINT is the occlusion of the camera's field of view by the embedded spiral channel. This limitation can be addressed by *image inpainting*, a well-studied task in computer vision that restores missing or occluded regions so that edges and textures plausibly continue across the hole. In practice, most inpainting pipelines expect a user-drawn binary mask to specify what to remove [26]–[28]. Yet accurate masks are burdensome and brittle, especially for thin, curvilinear occlusions like ours, since small misalignments or low contrast cause over-/under-inpainting and add per-frame manual overhead. Moreover, segmentation models often struggle when the occluder closely resembles the background, complicating automated mask generation. This mask dependency is increasingly avoidable: state-of-the-art text-guided editors can localize and modify regions from prompts alone while preserving surrounding structure. Approaches such as InstructPix2Pix (instruction-following editing) [29], Prompt-to-Prompt (attention-controlled, prompt-only localization) [30], Imagic and Null-Text Inversion (high-fidelity editing of real images) [31], [32], and SDEdit (robust diffusion-based refinements) [33] collectively demonstrate the practicality of segmentation-free, prompt-driven edits. In practice, modern production multimodal editors make these capabilities easy to use, enabling prompt-based local edits such as object removal and background cleanup. Motivated by these advances, in this work, as a proof-of-concept and without developing a complex computer vision pipeline, we adopt a segmentation-free, prompt-driven post-processing step using

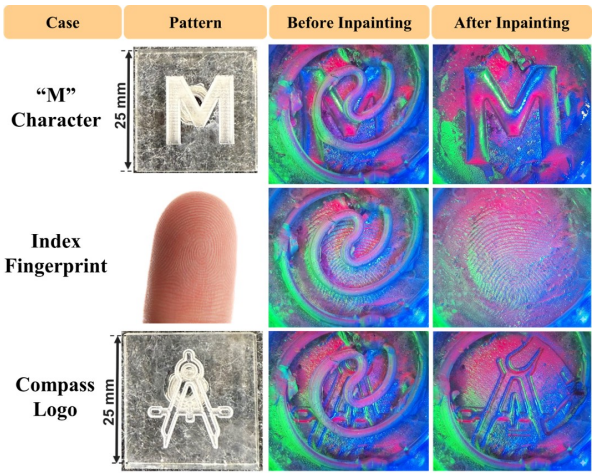

Fig. 5. Visualization of the inpainting result showing removal of the spiral channel and reconstruction of the underlying texture using a Gemini 2.5 flash image.

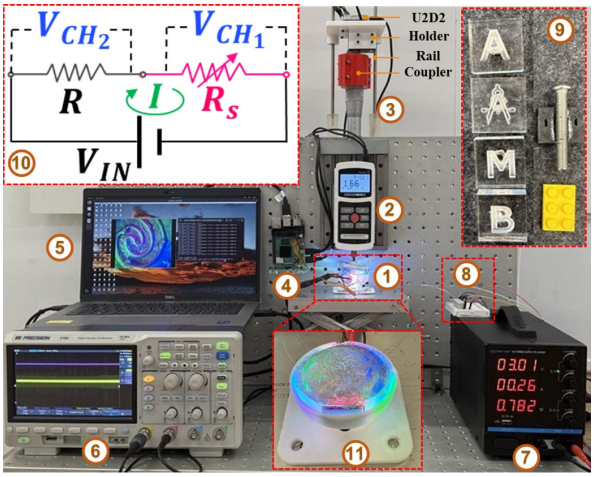

Fig. 6. Experimental Setup – The setup used for the characterization and calibration of the MINT sensor includes: (1) MINT sensor placed on a rigid housing, (2) Mark-10 Series 5 Digital Force Gauge for precise force measurement, (3) Motorized linear stage to provide precise vertical movement, (4) Raspberry Pi for processing the camera output, (5) PC for monitoring and controlling the experimental setup, (6) BK Precision 2194 Digital Storage Oscilloscope for measuring voltage changes in the circuit, (7) DC power supply to provide input voltage for the circuit, (8) Circuit board used for determining the resistance of spiral within the MINT sensor, (9) Various test objects used in the experiments, including different characters, compass logo, M6 screw, and Lego block 2×3, (10) Schematic view of the circuit for calculating the resistance of the MINT sensor, and (11) Zoomed-in view of MINT sensor placed on housing.

a state of the art, pretrained text-to-image generation and editing model: Gemini 2.5 Flash Image (aka *nano-banana*) [25]. We observe that the world understanding of this large-scale model is sufficient to zero-shot the inpainting process with only minor artifacts. The model is guided by a simple, direct natural language prompt to perform the inpainting task. A representative prompt is: "*Remove only the spiral artifact and estimate what is underneath*." The model, leveraging its extensive pre-training, identifies the spiral artifact based on the prompt's context and replaces it by generating a new texture that is coherent with the visible surrounding patterns. This prompt-based method requires no specialized model training, fine-tuning, or manual mask creation, offering a highly efficient and accessible solution to overcome the sensor's physical occlusion. As shown in Fig.5, its final output is a complete, high-fidelity texture map suitable for downstream analysis. To preserve the integrity of measurements, the edited frames are used only for visualization, and the original, unaltered frames are retained for all quantitative analyses.

III. Evaluation Experiments and Results

Before conducting a systematic analysis on fabricated sensors, we first performed preliminary analysis on the fabricated S1, S2, and S3 sensors (shown in Fig. 4) to evaluate their quality on texture mapping and maximum force measurement. After performing these experiments, S3 sensor was determined to be the optimal design among the three cases for further evaluation. As shown in Fig. 4, for sensor S1, the texture quality did not meet the required standard because the spiral channel obstructed the camera's field of view, limiting its ability to capture detailed textures. Additionally, the maximum allowable force that could be applied was too low and saturated fast before obtaining a minimally acceptable texture quality. To address this issue, we made sensor S2 in which the coil gap was increased to create a larger unobstructed window for capturing the texture of contacting objects. However, for this sensor, although the texture quality improved due to the adjusted coil gap, the issue of allowable force limitations persisted. The sensor was unable to measure forces beyond 2 N, and the response saturated quickly, as expected. Ultimately, Sensor S3 provided the best balance between texture capture quality and maximum measurable force, making it the most effective configuration for further testing and application. Considering this evaluation, we selected Sensor S3 for further systematic analysis.

A. Experimental Setup

Figure 6 illustrates the experimental setup used for the texture and force characterization and calibration of the fabricated S3 MINT sensor. As shown, this experimental setup is composed of the following components: A *Raspberry Pi* that functions as the microcontroller responsible for capturing and processing video output from the camera to observe texture and stream it to a display. For precise force measurements, *a digital force gauge (Mark-10 Series 5, Mark-10 Corporation)* with 0.02 N resolution is used as a reference for calibration. The *motorized linear stage* (M-UMR12.40, Newport), offering 1 μm precision, enables highly accurate movement of the force gauge probe based on linear stage and motor encoder data. To automate movement, *a custom-designed mechanism* integrates a DC motor (Dynamixel XM430-W350-R), coupling, and nickel-plated chrome steel rods (6 mm) with linear ball bearings (LM6UU) to efficiently transfer motor torque to the micro-stage. A *U2D2 controller and U2D2 power hub* are used to control the motor and read the 12-bit encoder data from the DC motor. A *digital storage oscilloscope (BK Precision 2194, B&K Precision)* with 0.1 mV resolution measures voltage changes across resistors in the circuit, enabling the calculation of spiral channel resistance within the MINT sensor. The *DC power supply* (LWLONGWEI PS-

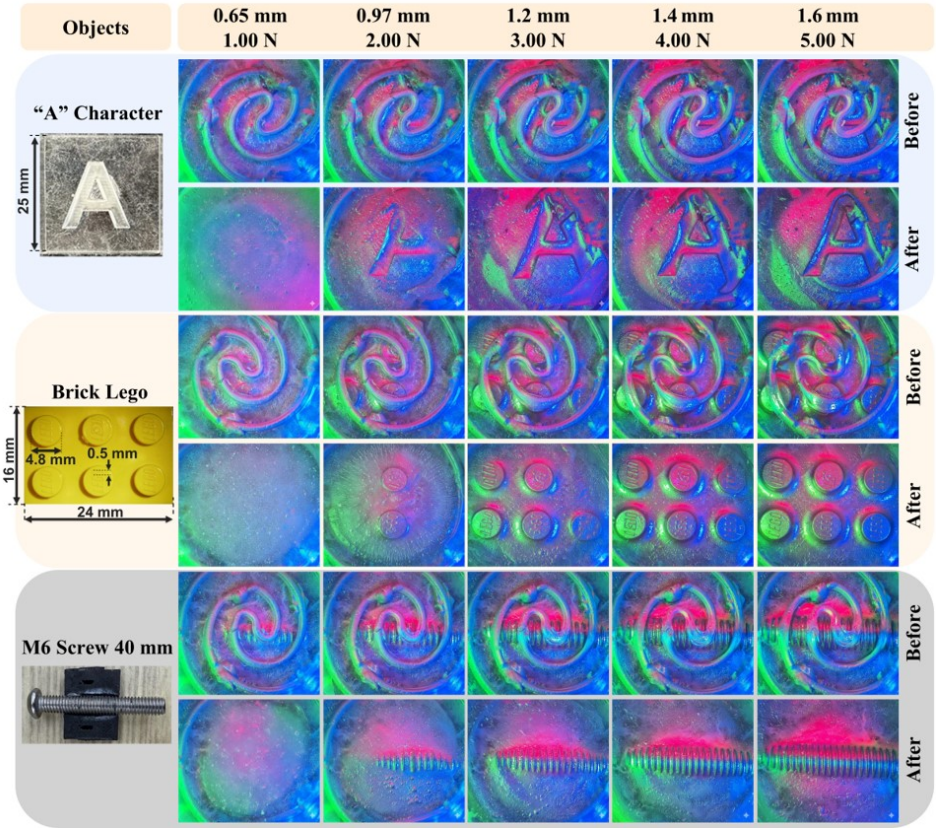

Fig. 7. Visual outputs of the MINT Framework before and after inpainting during interactions with various objects, including "A" character, a Lego block 3×2, and an M6 screw with 40 mm length, captured at different interaction force levels ranging from 0 N to 5 N.

3010DF) provides the input voltage required for determining the unknown and variable resistance of the spiral channel. A series circuit is used to determine the resistance of the spiral channel. This circuit consists of a $V_{IN} = 3V$ DC power supply, a fixed resistor with a nominal resistance of 10 Ω (actual measured value: $R = 9.5\Omega$), and a variable spiral resistor embedded within the MINT sensor (R_s). Of note, the resistance of the unknown spiral resistor within the gel is calculated by measuring the voltage across both the variable spiral resistor and the fixed resistor using Channel 1 (CH_1) and Channel 2 (CH_2) of the oscilloscope, respectively. Then, by applying Kirchhoff's Law, the resistance of the spiral channel is determined as follows:

$$\begin{cases} I = \frac{V_{CH_1}}{R} \\ I = \frac{V_{CH_2}}{R_s} \end{cases} \Rightarrow R_s = \frac{V_{CH_1}}{V_{CH_2}} R \tag{1}$$

Of note, since the sum of voltages across channels 1 and 2 equals the fixed voltage of the DC power supply, the resistance of the spiral channel embedded within the substrate can be calculated by measuring only the voltage across channel 1 or 2, as follows:

$$R_s = \frac{V_{CH_1}}{V_{IN} - V_{CH_1}} R = \frac{V_{IN} - V_{CH_2}}{V_{CH_2}} R \tag{2}$$

However, due to the internal resistance of the power supply and the relatively small magnitude of the spiral channel resistance, (1) was used to reduce calculation errors. Of note, the schematic circuit shown in Fig. 6 represents the complete setup used to determine the resistance of the spiral channel embedded in the substrate.

B. Experimental Procedure

To conduct the experiments and evaluate the texture mapping and force measurement of the S3 sensor, we first hold this sensor on the rigid frame. Next, as shown in Fig. 6, various objects with different shape and texture patterns were attached to the force gauge. These objects included (i) a laser cut "A", "B", and "M" characters, (ii) a laser cut Compass logo, (iii) Lego block $2mm \times 3mm$, and (iv) an M6 screw. After the hardware was set up for the measurements, using a custom-code written in the python, camera, oscilloscope, force gauge, and linear stage motor were initialized together for seamless coordination among all components, ensuring accurate and synchronized data acquisition throughout the experiment. Once the desired movement (i.e., 1/16 mm) was achieved through the linear actuation system, the system paused for 10 seconds. Then, a customized program was activated to collect data at one-second intervals. This process was repeated until a force of 5 N was reached, at which point the data collection was halted. Throughout the process, a custom-written Python script controls data collection within the Robot Operating System (ROS) environment, enabling full automation of the experimental procedure. This minimizes human-related errors and ensures a consistent and repeatable process. This process includes streaming and recording the camera outputs, calculating linear stage movement based on encoder data,

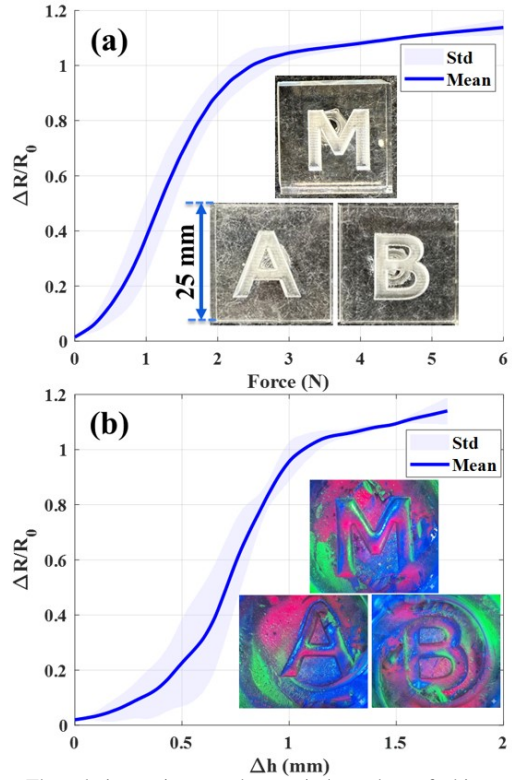

Fig. 8. The relative resistance change, independent of object textures, is shown as a function of (a) the applied force measured by the force gauge and (b) the vertical displacement of the linear stage. The results are reported as the mean value (solid line) with the standard deviation (shaded area) across three different cases. In addition, the captured image after inpainting at an applied force of 5 N is included.

force reading, and calculating resistance based on all voltage measurements by oscilloscope.

Figures 7 illustrates the textural outputs of S3 sensor before and after inpainting during interactions with three representative objects (i.e., a Lego block, a laser cut A character, and an M6 screw) at different interaction force (i.e., 0 to 5 N). As can be observed, with increasing the interaction force, the captured texture details became more refined, demonstrating the sensor's ability to enhance texture resolution with increasing force. Further, to establish the relationship between the variable resistance of the spiral channel (R_s) and the normal interaction force (F_N) through a calibration procedure, we used laser cut objects (i.e., "A", "B", and "M" characters) and followed the experimental procedure described above. As shown in Fig. 8, these objects have the same contact surface, which is crucial for correctly calibrating the MINT sensor and finding the relationship between the interaction force and measured resistances of the spiral channel. Figure 8 exhibits the relationship between the relative resistance change and the applied force and vertical displacement, respectively.

IV. Discussion and Conclusion

The textural images after inpainting, shown in Fig.7, and the measured forces, shown in Fig.8, demonstrate that the fabricated S3 MINT framework, compared to conventional VTS, can not only capture texture but also directly measure interaction force. This capability of the MINT framework allows real-time force measurement without relying heavily on the ML and AI algorithms used in [9], [34]. In particular, Fig. 5 and Fig. 7 clearly indicate ability of MINT framework in capturing fine details of a texture such as human fingerprint and the word "*LEGO*" on a Lego block using the prompt-based inpainting method. Also, investigation of Fig. 7 indicates the sensitivity of S3 sensor in capturing fine details of the objects in low interaction forces as small as 3 N. Of note, as shown in the progressive images in Fig. 7, the prompt-based inpainting approach reveals increasingly finer details as the interaction force increases. These results clearly demonstrate the strong performance of the prompt-based inpainting approach using Gemini 2.5 Flash Image (aka nano-banana) [25]. The model's broad world understanding enables it to perform zero-shot inpainting with only minor artifacts, guided by a simple and direct natural language prompt. This approach requires no specialized model training, fine-tuning, or manual mask creation, providing an efficient and accessible solution to address the sensor's physical occlusion.

Also, Fig. 8 illustrates a sensitive and approximately linear relationship between the measured resistance of the spiral channel and the interaction forces for forces below 3 N. Beyond this threshold, the sensor's sensitivity (i.e., the slope of the plot) decreases, leading to a separate linear relationship. Additionally, irrespective of laser-cut blocks with varying textures, the S3 sensor demonstrates consistent performance, highlighting the repeatability and robustness of the calibration process. Furthermore, Fig. 8 reveals the presence of a small dead zone for forces below 0.25 N, which can be attributed to the soft and deformable nature of MINT's flexible gel layer. This layer absorbs a portion of the applied interaction force, resulting in minimal or negligible resistance changes within this range. Notably, the dead zone is directly correlated with the small depth of the spiral channel relative to the sensor's surface (i.e., 0.5 mm); increasing this depth reduces the sensitivity of force measurements and consequently enlarges the dead zone.

V. Conclusion and Future Directions

In this study, we introduced design, fabrication and evaluationof MINT as a novel vision-based Soft Sensor for direct measurement of force and texture. This innovative approach allows for more accurate and safer interactions with delicate objects, making it readily adaptable for various devices and applications. Through performing various experiments, we demonstrated that MINT framework is able to capture fine textural details of various objects at forces greater than 3 N. We also obtained a piecewise linear and repeatable relationship between the measured resistance of the embedded spiral channel and the applied interaction forces demonstrating the high sensitivity and robustness of proposed fabrication procedure.

In the future, we plan to integrate the artificial intelligence techniques [35] with MINT sensor in a real-time control loop, allowing the system to automatically adjust the interaction force based on live texture capture. This would

make the MINT framework even more versatile and adaptive, providing an optimal balance between force measurement and texture acquisition for complex tasks such as pipe inspection using soft robots [36], [37]. We will also use this proposed framework for a more realistic application such as robotic object manipulation by integrating the sensor with a gripper and robotic arms [38].

REFERENCES

[1] S. Zhang, Y. Yang, Y. Sun, N. Liu, F. Sun, and B. Fang, "Artificial skin based on visuo-tactile sensing for 3d shape reconstruction: Material, method, and evaluation," *Advanced Functional Materials*, 2024.

[2] M. K. Johnson and E. H. Adelson, "Retrographic sensing for the measurement of surface texture and shape," *2009 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1070–1077, 2009.

[3] O. C. Kara, N. Ikoma, and F. Alambeigi, "Hysense: A hyper-sensitive and high-fidelity vision-based tactile sensor," *2022 IEEE Sensors*, pp. 1–4, 2022.

[4] H. Kim, O. C. Kara, and F. Alambeigi, "A soft and inflatable vision-based tactile sensor for inspection of constrained and confined spaces," *IEEE Sensors Journal*, vol. 23, no. 23, pp. 29 605–29 618, 2023.

[5] O. C. Kara, N. Venkatayogi, N. Ikoma, and F. Alambeigi, "A reliable and sensitive framework for simultaneous type and stage detection of colorectal cancer polyps," *Annals of Biomedical Engineering*, pp. 1–14, 2023.

[6] B. Ward-Cherrier, N. Pestell, L. Cramphorn, B. Winstone, M. E. Giannaccini, J. M. Rossiter, and N. F. Lepora, "The tactip family: Soft optical tactile sensors with 3d-printed biomimetic morphologies," *Soft Robotics*, vol. 5, pp. 216 – 227, 2018.

[7] C. Sferrazza and R. D'Andrea, "Design, motivation and evaluation of a full-resolution optical tactile sensor," *Sensors (Basel, Switzerland)*, vol. 19, 2019.

[8] W. Yuan, S. Dong, and E. H. Adelson, "Gelsight: High-resolution robot tactile sensors for estimating geometry and force," *Sensors (Basel, Switzerland)*, vol. 17, 2017.

[9] A. Sipos, W. van den Bogert, and N. Fazeli, "Gelslim 4.0: Focusing on touch and reproducibility," *ArXiv*, vol. abs/2409.19770, 2024.

[10] W. K. Do, B. Jurewicz, and M. Kennedy, "Densetact 2.0: Optical tactile sensor for shape and force reconstruction," *2023 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 12 549–12 555, 2022.

[11] D. Zhang, W. Fan, J. Lin, H. Li, Q. Cong, W. Liu, N. F. Lepora, and S. Luo, "Design and benchmarking of a multimodality sensor for robotic manipulation with gan-based cross-modality interpretation," *IEEE Transactions on Robotics*, vol. 41, pp. 1278–1295, 2025.

[12] M. Lambeta, P. wei Chou, S. Tian, B. Yang, B. Maloon, V. R. Most, D. Stroud, R. Santos, A. Byagowi, G. Kammerer, D. Jayaraman, and R. Calandra, "Digit: A novel design for a low-cost compact high-resolution tactile sensor with application to in-hand manipulation," *IEEE Robotics and Automation Letters*, vol. 5, pp. 3838–3845, 2020.

[13] P. Yu, W. ting Liu, C. Gu, X. Cheng, and X. Fu, "Flexible piezoelectric tactile sensor array for dynamic three-axis force measurement," *Sensors (Basel, Switzerland)*, vol. 16, 2016.

[14] Q. Shu, Y. Pang, Q. Li, Y. Gu, Z. Liu, B. Liu, J. Li, and Y. Li, "Flexible resistive tactile pressure sensors," *Journal of Materials Chemistry A*, 2024.

[15] Z. Li, K. Zhao, J. Wang, B. Wang, J. Lu, B. Jia, T. Ji, X. Han, G. Luo, Y. Yu, L. Wang, M. Li, Z. Wang, and L. Zhao, "Sensitive, robust, wide-range, and high-consistency capacitive tactile sensors with ordered porous dielectric microstructures." *ACS applied materials & interfaces*, 2024.

[16] W. Peng, R. Zhu, Q. Ni, J. Zhao, X. Zhu, Q. Mei, C. Zhang, and L. Liao, "Functional tactile sensor based on arrayed triboelectric nanogenerators," *Advanced Energy Materials*, 2024.

[17] N. Yao and S. Wang, "Recent progress of optical tactile sensors: A review," *Optics & Laser Technology*, 2024.

[18] C. Wang, C. Liu, F. Shang, S. Niu, L. Ke, N. Zhang, B. Ma, R. Li, X. Sun, and S. Zhang, "Tactile sensing technology in bionic skin: A review." *Biosensors & bioelectronics*, vol. 220, p. 114882, 2022.

[19] M. Negre, M. Jorda, O. Vardoulis, A. Chortos, O. Khatib, and Z. Bao, "A hierarchically patterned, bioinspired e-skin able to detect the direction of applied pressure for robotics," *Science Robotics*, vol. 3, 2018.

[20] Y.-L. Park, B.-R. Chen, and R. J. Wood, "Design and fabrication of soft artificial skin using embedded microchannels and liquid conductors," *IEEE Sensors journal*, vol. 12, no. 8, pp. 2711–2718, 2012.

[21] Y. Mengüç, Y.-L. Park, E. Martinez-Villalpando, P. Aubin, M. Zisook, L. Stirling, R. J. Wood, and C. J. Walsh, "Soft wearable motion sensing suit for lower limb biomechanics measurements," in *2013 IEEE International Conference on Robotics and Automation*. IEEE, 2013, pp. 5309–5316.

[22] Y.-L. Park, C. Majidi, R. Kramer, P. Bérard, and R. J. Wood, "Hyperelastic pressure sensing with a liquid-embedded elastomer," *Journal of micromechanics and microengineering*, vol. 20, no. 12, p. 125029, 2010.

[23] M. R. Javazm, O. C. Kara, and F. Alambeigi, "A novel soft and inflatable strain-based tactile sensing balloon for enhanced diagnosis of colorectal cancer polyps via colonoscopy," *IEEE Sensors Journal*, vol. 24, pp. 26 564–26 573, 2024.

[24] M. R. Javazm, S. Kiehler, O. C. Kara, and F. Alambeigi, "Towards design and development of a soft pressure sensing sleeve for performing safe colonoscopic procedures," *2024 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 12 787–12 792, 2024.

[25] Google DeepMind. (2025) Gemini 2.5 flash image. [Online]. Available: https://developers.googleblog.com/en/introducing-gemini-2-5-flash-image/

[26] L. Zhang, A. Rao, and M. Agrawala, "Adding conditional control to text-to-image diffusion models," in *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, 2023.

[27] G. Liu, F. A. Reda, K. J. Shih, T.-C. Wang, A. Tao, and B. Catanzaro, "Image inpainting for irregular holes using partial convolutions," in *Proceedings of the European Conference on Computer Vision (ECCV)*, September 2018.

[28] R. Suvorov, E. Logacheva, A. Mashikhin, A. Remizova, A. Ashukha, A. Silvestrov, N. Kong, H. Goka, K. Park, and V. Lempitsky, "Resolution-robust large mask inpainting with fourier convolutions," in *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*, 2022.

[29] T. Brooks, A. Holynski, and A. A. Efros, "Instructpix2pix: Learning to follow image editing instructions," *arXiv preprint arXiv:2211.09800*, 2022.

[30] A. Hertz, R. Mokady, J. Tenenbaum, K. Aberman, Y. Pritch, and D. Cohen-Or, "Prompt-to-prompt image editing with cross attention control," *arXiv preprint arXiv:2208.01626*, 2022.

[31] B. Kawar, S. Zada, O. Lang, O. Tov, H. Chang, T. Dekel, I. Mosseri, and M. Irani, "Imagic: Text-based real image editing with diffusion models," *arXiv preprint arXiv:2210.09276*, 2022.

[32] R. Mokady, A. Hertz, K. Aberman, Y. Pritch, and D. Cohen-Or, "Null-text inversion for editing real images using guided diffusion models," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2023.

[33] C. Meng, Y. He, Y. Song, J. Song, J. Wu, J. Zhu, and S. Ermon, "Sdedit: Guided image synthesis and editing with stochastic differential equations," in *International Conference on Learning Representations (ICLR)*, 2022.

[34] I. H. Taylor, S. Dong, and A. Rodriguez, "Gelslim 3.0: High-resolution measurement of shape, force and slip in a compact tactile-sensing finger," *2022 International Conference on Robotics and Automation (ICRA)*, pp. 10 781–10 787, 2021.

[35] S. Kapuria, N. Ikoma, S. P. Chinchali, and F. Alambeigi, "Synthetic data-augmented explainable vision transformer for colorectal cancer diagnosis via surface tactile imaging," *Eng. Appl. Artif. Intell.*, vol. 151, p. 110633, 2025.

[36] L. V. Nguyen, H. Kim, K. T. Nguyen, F. Alambeigi, and V. A. Ho, "Adaptable cavity exploration: Bioinspired vibration-propelled pufferface robot with a morphable body," *Science Advances*, vol. 11, 2025.

[37] M. R. Javazm, S. Kiehler, V. A. Ho, and F. Alambeigi, "Design and characterization of a vibration-driven pufferfish-inspired inflatable soft robot for pipe inspection," *International Conference on Soft Robotics (RoboSoft)*, vol. 6, 2026.

[38] O. C. Kara, M. R. Javazm, O. Rezayof, and F. Alambeigi, "Intact-grip: An inflatable tactile gripper for soft manipulation and high-resolution texture mapping," in *2026 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2026.