Deep learning-based tactile sensor design and interpretability of image recognition

Qingling Duan*, Qi Zhang*, Zhiyuan Liu and Yongsheng Ou

Abstract— **Good tactile feedback (such as contact force) can provide rich proprioception in complex tactile contact task scenarios such as robot dexterous operation. However, tactile sensors applied in robotics lack accurate 3D force decoupling. Here, we present a soft tactile sensor with self-decoupling by designing colored silicone blocks, whose deformation can be detected by a camera according to the change of the image of the colored silicone block under external forces. Furthermore, by designing different flexible probe pattern layers, the connection between the flexible probe pattern and the deep learning image recognition mechanism is studied to guide the design of the flexible sensor probe. As a result, the sensor can accurately measure the normal and shear forces through deep learning. Furthermore, by mounting our sensor at the fingertip of a robotic gripper, the robot can perform challenging tasks, such as grasping vulnerable objects and measuring force-change curves during the water-adding process. This research provides new insight into tactile sensor design and could be beneficial to various applications in the robotics field.**

Index Terms— **Deep learning in robotics and automation, force and tactile sensing, force-decoupling mechanism of the images**

I. INTRODUCTION

With the rapid development of computer vision, robots' visual perception and understanding of unstructured and natural scenes have rapidly improved [1]. However, in contact-rich task scenarios such as complex and dexterous robot operation, tactile perception often obtains more direct, accurate and rich proprioception than vision, thus generating a more reliable operation and control strategy [2, 3]. However, over the years, despite significant progress in robot operations [4], achieving good tactile feedback (e. g., contact force) and dexterous daily operations (e. g., adaptive grasping) remains a significant challenge. One of the main reasons is that robots lack soft tactile sensing systems that accurately perceive subtle changes in forces. Recently, rigid tactile sensors, such as force-sensitive resistance, lack soft, deformable surfaces that facilitate physical environment interactions. Therefore, designing the soft force sensor is critical for the robot field, which can solve the current difficulties facing robots and promote the development of robots.

Traditional tactile sensors rely mainly on electrical [5, 6 ,7, 8] or magnetic [9, 10, 11] modes. However, they are prone to interference and failure in electromagnetic fields, limiting their application scenarios [12, 13]. Optical-based tactile sensors [14, 15, 16] effectively avoid this problem and provide greater precision, sensitivity, and reproducibility [17, 18]. Among these sensors, the vision-based tactile image sensors [19, 20] can express complex tactile information directly on the images captured by the camera. Meanwhile, deep learning has powerful image processing capabilities [21, 22], which is an effective method for extracting complex tactile data from images [23]. However, optical-based tactile sensors are usually bulky [14] and unsuitable for daily grasping and operation tasks [24]. Moreover, deep learning algorithms are generally not highly interpretable, significantly reducing the reliability of deep learning-based schemes, which is not conducive to developing deep learning in applications.

To address the mentioned challenges, this paper designs a soft tactile sensor with force sensing capability and three-dimensional force self-decoupling capability, which can be adapted for daily robot grasping and operation tasks. The sensor consists of coloured silicone, a camera, and light sources. When the sensor's surface is forced, the silicone will deform, causing the colour and intensity of the light emitted by the light source to be reflected by the silicon, and the pattern captured by the camera will change. Furthermore, this paper designs different flexible probe pattern layers to study the image recognition mechanism (interpretability) of deep learning in the process of 3D force decoupling to guide the design of the flexible sensor probe. By climbing tactile sensors at its fingertips, the robot can use tactile feedback for challenging tasks, such as stably grasping vulnerable objects and measuring force change curves during the water process.

II. RELATED WORK

The traditional soft force sensors generally detect one-dimensional force (tension or pressure) through the changes in capacitance [5, 6] and resistance [7, 8] caused by the material deformation of external force. However, this method is challenging to detect three-dimensional force. Because in the 3D force detection process, the shear force(F_x, F_y) in

This work was supported in part by the Key-Area Research and Development Program of Guangdong Province (2019B090915002); in part by the National Natural Science Foundation of China (Grants No. U1813208, 62173319,62063006); in part by the Guangdong Basic and Applied Basic Research Foundation (2020B1515120054) and in part by the Shenzhen Fundamental Research Program (JCYJ20200109115610172). (Corresponding author: Yongsheng Ou.)

Qingling Duan is with Faculty of Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences, Shenzhen 518055, Guangdong Province, P.R. China, and also with the Department of Shenzhen College of Advanced Technology, University of Chinese Academy of Sciences, Shenzhen, China.

Qi Zhang and Zhiyuan Liu are with Faculty of Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences, Shenzhen 518055, Guangdong Province, P.R. China.

Yongsheng Ou is with Faculty of Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences, Shenzhen 518055, Guangdong Province, P.R. China, and also with the Department of Guangdong Provincial Key Lab of Robotics and Intelligent System, Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences. (e-mail: ys.ou@siat.ac.cn)

*These authors contributed equally to this work.

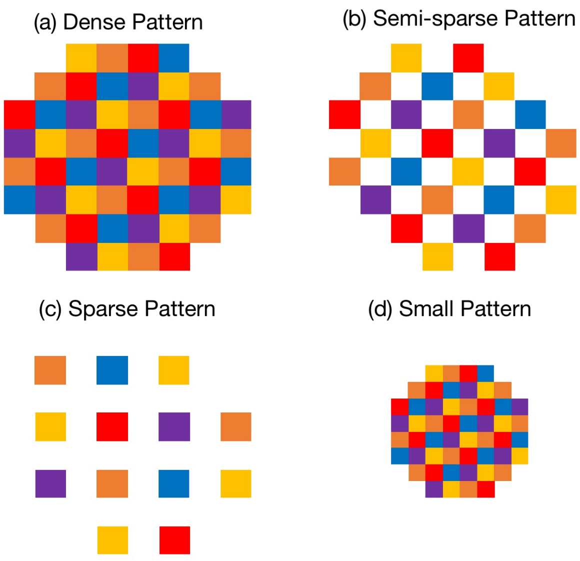


Fig. 1. Tactile Image Sensor Pattern Designs

the plane direction and the normal force(F_z) in the vertical direction will simultaneously cause the deformation of the sensor, making the resulting signals interfere with each other [25]. Therefore, researchers have tried to decouple three-dimensional forces under soft conditions through mathematical modelling [26, 27], structural innovation [28, 29], and other methods. For example, in [30], Songyue Chen et al. reported a flexible three-dimensional force sensor based on the hemisphere interlocked structure, which measures the normal force or shear force separately through the contact resistance changes of the three interlocked hemisphere units, and still faces the problems of complex decoupling process, easy decoupling interference and low decoupling accuracy.

With the development of image processing technology, it has become a new idea to transform multidimensional force into an image signal and decouple multidimensional force through image processing. This idea was implemented in the visual-tactile sensor Gelsight, proposed by MIT's E. H. Adelson team in 2014 [31]. They introduced marker points on the reflective film inside the soft elastomer to capture the displacement of marker points under the 3D force. Furthermore, they established the mapping relationship between the marker point displacement and 3D force through finite element analysis to realize 3D force detection in a soft environment. However, the established three-dimensional force model has a small application range due to the need to set many preconditions and simplify the problem.

Deep learning has become a mainstream method for handling familiar objects in many necessary fields, such as text [32], image [21], video [33], and graphics [34]. Therefore, combining deep learning with multidimensional force decoupling becomes a new option. For example, Baimukashev et al. [35] proposed a tactile image sensor combining an optical fibre and a camera that successfully decoupled the three-dimensional forces with a multi-output CNN model. Kakani et al. [36] improved on the VGG-16 deep neural network to realize the measurement of the contact location, the contact region, and the contact force distribution of the binocular tactile images. Yuan et al. [37] performed measurements of 3D forces and normal moments by directly inputting GelSight tactile images into neural networks.

However, the mechanism of deep learning is imperfect, resulting in deep learning still being in the "black box" state. For example, in [38], D. Heaven points out that without understanding the deep learning recognition mechanism, the small changes in the images may seriously impact the results, which significantly reduces the reliability of deep learning-based solutions and is not conducive to developing deep learning in engineering applications. At the same time, because the identification mechanism of deep learning was not understood in general, deep learning is only an auxiliary tool for data processing in most applications, so it cannot provide practical guidance for designing sensors.

Therefore, studying the image recognition mechanism is crucial for applying deep learning in 3D force decoupling, which is related to the application's stability and the sensor's reliability, but also a guide for the sensor probe design and algorithm improvement. Therefore, this paper studies the connection between flexible probe patterns and deep learning mechanisms by designing different flexible probe pattern layers, as shown in Fig. 1. Furthermore, this paper studies the deep learning mechanism through deep learning interpretability technology [39] to guide the front-end sensor design. The contribution of this study is summarized as follows: 1)Proposed a sensor preparation process and proposed the 3D deep learning-based force decoupling method under soft conditions. 2)Studied the image recognition mechanism of different soft probe patterns in deep learning of 3D force decoupling and explored the sensor design through deep learning mechanism. 3)Realized the soft 3D force sensor preparation to reach the engineering application level. The sensor can be fixed to the fingertips of the robot gripper to grab fragile objects.

III. Design of Tactile Image Sensor

A. Fabrication Method

The flexible probe of the sensor is mainly made of silicone and dye, shaped by mold. For selecting materials, the flexible probe is the central part of the three-dimensional force sensor, which must withstand thousands of extrusions during data acquisition and therefore has high requirements for material softness, durability and tear resistance. At the same time, in order to be able to collect a clearer 3D map image, the material also needs to have better transparency. The Sorta-clear^{TM} 12 silicone gel from Smooth-on has excellent softness(Shore hardness 12A), stretchability (maximum elongation 590%), and tear-resistance(11558 N/m), high transparency, and matching pigments that can be used to prepare pattern layers. In this paper, Sorta-clear^{TM} 12 silicone was mixed with different pigments to make the base and pattern layers, fixed the pattern layer through the Sil-Poxy^{TM} silicone glue of Smooth-on, and then filled the

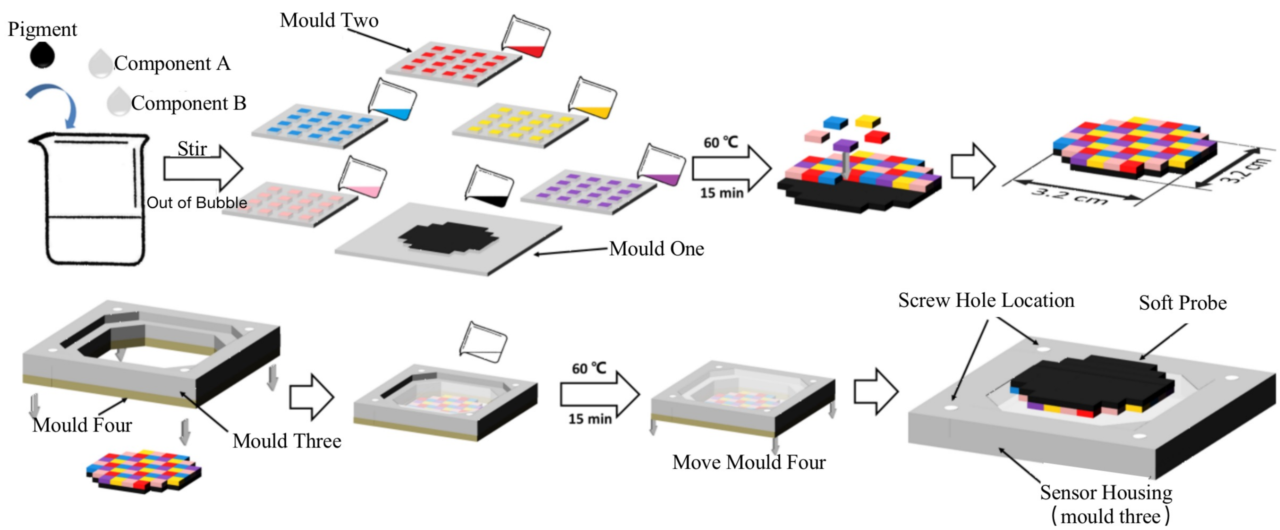


Fig. 2. The Soft Probe Preparation Process

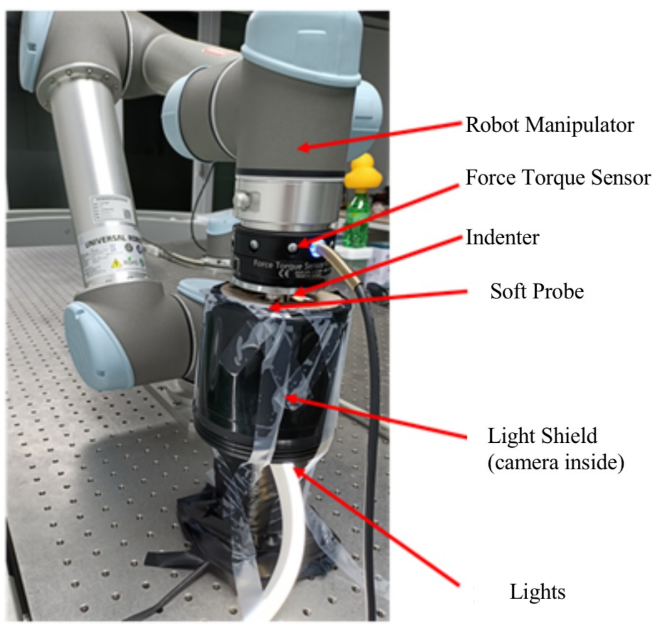


Fig. 3. Data Acquisition Platform

pure Sorta-clear^{TM} 12 silicone as the transparent layer. The specific study method is shown in Fig. 2 and is described as follows:

1) Mix Sorta-clear^{TM}12 silicone according to mass ratio A: B=1:1, add black pigment of 5% of total silicone quality and rotate it into the deflating machine for 4 minutes. The black liquid silicone is poured into mould one and heated 60°C for 15 minutes to obtain the base layer;
2) In the same ratio as the step(1), pour the mould 2 to prepare the pattern layer silicone, get small yellow, red, blue, purple and pink silicone of 4 mm×4mm, and fix them in the base layer by Sil-Poxy^{TM} silicone;
3) Fix the moulds 3 and 4 around the pattern layer, pour 1:1 Sorta-clear^{TM}12 silicone, heat and cure to form a transparent layer. After the preparation of the soft probe is completed, mould four is removed, and the mold 3 can serve as the shell of the soft probe and is fixed on the automatic acquisition platform with screws through the reserved hole position.

B. Pattern Principle

The main idea of decoupling the 3D force is to record the deformation of the sensor through the camera. For this type of visual-tactile sensor, the pattern is integral to its expressed information. In order to observe the impact of patterns on neural network output, this paper designed four probes with different degrees of complexity, depending on two factors: the density and size of silicone blocks. All but the different patterns are made from the same material. Among these, patterns Fig. 1(a), Fig. 1(b) and Fig. 1(c) differ in density, with Fig. 1(a) being the densest, Fig. 1(c) being the most sparse, and B being somewhere between them. Fig. 1(a) and Fig. 1(d) are used to contrast the effects of squares of different sizes. The sparseness and miniaturization of patterns can help reduce the complexity of device production and accelerate the mass production of sensors.

IV. Experimental Platform

A. System Integration

As is shown in Fig. 3, in addition to the prepared elastomer, the experimental system includes the light source, camera, force sensor, and mechanical arm.

- Light source. It is recommended to use 3w led lamp beads(Bridgelux, USA, cob lamp bead) as the light

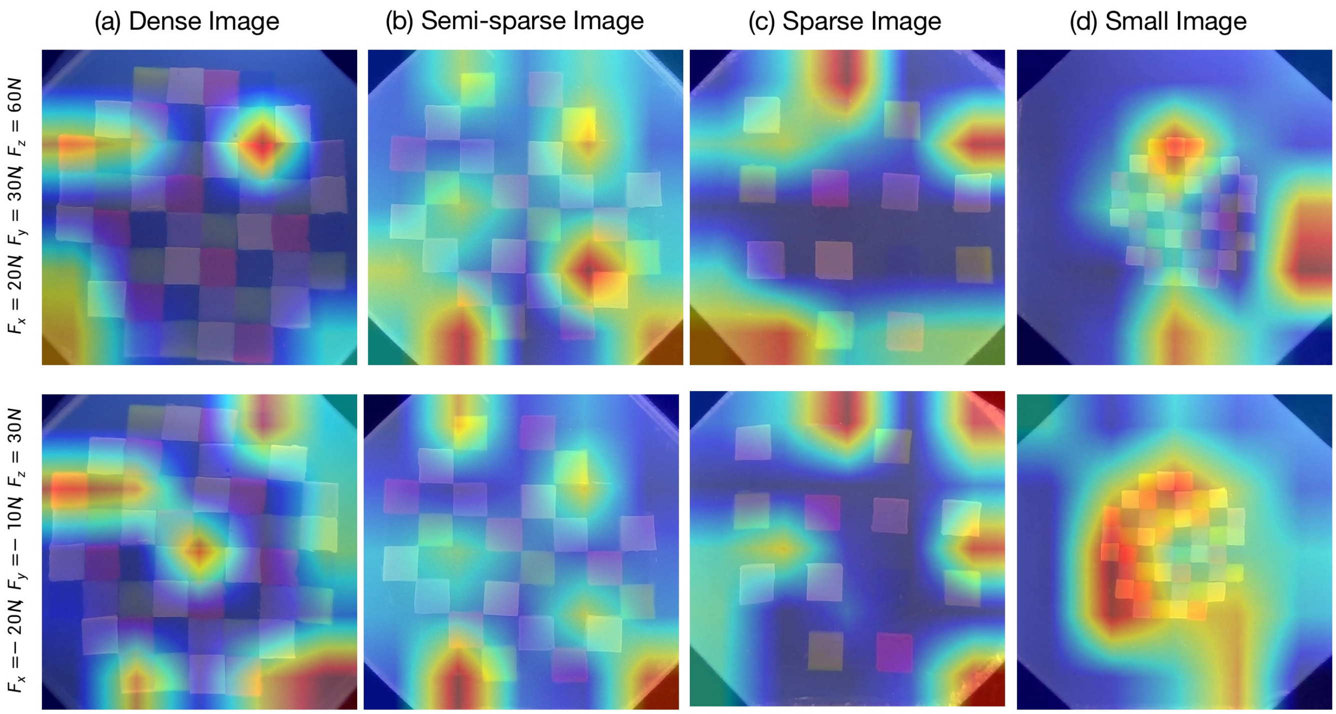


Fig. 4. The CAM Visualized using the Grad-CAM Method

source and avoid using colored light or other special light. Secure the LED lamp to the heat sink to avoid overheating.

- Camera. The camera used is a CCD industrial camera with a resolution of 819×819, and the camera is placed at the bottom of the sensor to capture the deformation pattern. Communicate with the computer via USB.
- Force sensor. A commercially available torque sensor(ROBOTIC, FT 300-S Force Torque Sensor) is fixed to the end of the robotic arm, which is used to record the force applied by the robotic arm.
- Mechanical arm. The main body of the collection platform is an industrial robot(UNIVERSAL ROBOTS, UR5e Robot), which is used as force application equipment, and its end is equipped with a torque sensor. In addition, a 20mm diameter pressure head is installed at the end of the torque sensor, which acts on the soft probe surface.

B. Experimental Operation

The data acquisition program was written in Labview, which was realized to control the robot manipulator to move to the specified position, automatically obtain the camera image, and synchronously record the 3-dimensional force information. Communication with the mechanical arm uses the TCP/IP protocol. A position servo of 125hz was used to control the arm-end-indenter applied to the soft probe at the same speed. The fabricated sensor is fixed to the optical platform. Before each experiment, the robotic arm's end was moved to the top of the sensor, serving as the start point. The data acquisition process is to reach a given eight depths at a speed of 2mm/s, form different normal forces, under each depth to 1000 different positions, forming a different shear force, back to the starting point after each shear force is applied. Finally, 24000 ($1000 \times 8 \times 3$) group sampled data was obtained. The resulting dataset is divided into training sets (70%) and validation sets (30%) for the training and validating parts of the deep learning model.

TABLE I

AN EXAMPLE OF A TABLE

	Pattern A	Pattern B	Pattern C	Pattern D
F_x/N	0.41	**0.26**	0.35	0.37
F_y/N	0.45	**0.27**	0.41	0.43
F_x/N	0.66	**0.45**	0.7	0.66

Note: The error is calculated RMSE error

V. SENSOR CHARACTERIZATION

A. The Connection between Soft Probe Pattern and Deep Learning Mechanism

The structure of the optic-tactile sensor, such as the type of flexible material, pattern, production process, Etc., largely determines its performance. Therefore, this paper mainly discusses the influence of patterns on the resolution of tactile sensors, keeping the experimental platform, the number of datasets and the deep learning model consistent except for the probe pattern changes.

Here, this paper analyze the error (RMSE) of the decoupled three-dimensional forces (F_x, F_y, F_z) under four

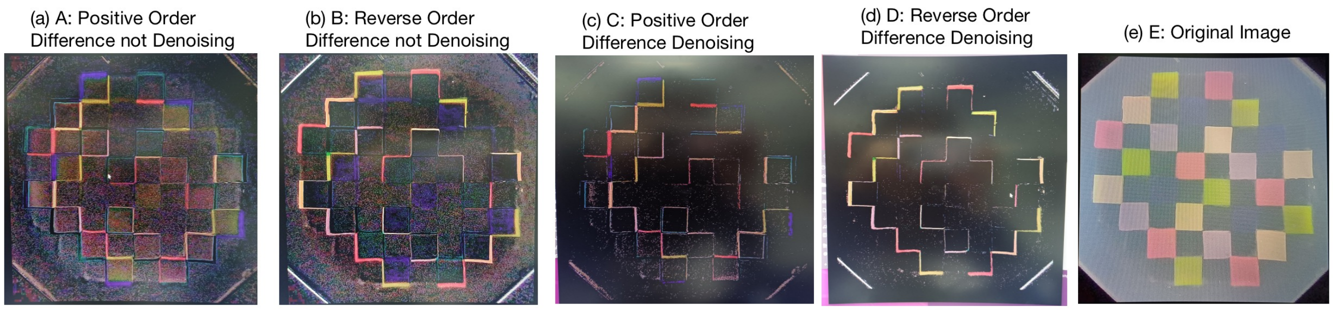
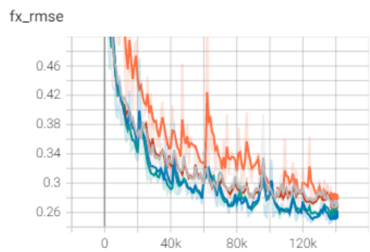
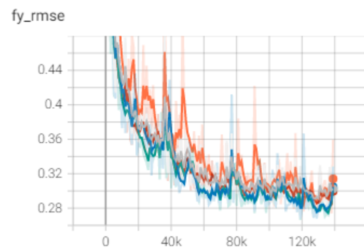


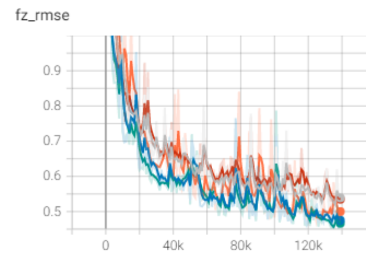
Fig. 5. Five Different Preprocessing Combinations



(a) F_x result



(b) F_y result



(c) F_z result

Fig. 6. The evaluation of the five different preprocessing collected by the semi-sparse pattern: grey line: preprocessing A, blue line: preprocessing B, red line: preprocessing C, green line: preprocessing D, orange line: preprocessing E

different patterns. The errors in the test set under different patterns are shown in Table 1. Among them, the network trained on the data collected by the semi-sparse pattern performs best on the validation set, where the RMSE of F_x error is 0.26N, the RMSE of F_y is 0.27N, the RMSE of F_z is 0.45N. Other motifs have a F_x error between 0.38N $\pm$ 0.03N, a F_y error between 0.43N $\pm$ 0.02N, a F_z error between and 0.68N $\pm$ 0.02N. It is shown that the proper sparsity helps to improve the accuracy. However, extreme sparsity decreases the accuracy. Reducing the size of the color block has less effect on the accuracy.

To better understand the reasons for this result, to visualize the CAM using the Grad-CAM method [40], which is used to locate the sensitive regions of the neural network model. In the visualization example of Fig. 4, stronger CAM regions used brighter colours. This paper compared the performance of the four patterns on (F_x = 20N, F_y = 30N, F_z = 60N) and (F_x = -20N, F_y = -10N, F_z = 30N). Due to the influence of the light source, it tends to cover the edge part (sunny position), which is obvious in the small pattern. Because the semi-sparse pattern performs better, it tends to cover the entire image. This ability to accurately locate the stress region in the CAM map species has a potential value for the image decoupling force.

B. Effect of Different Preprocessing Methods on Image Recognition

The semi-sparse pattern was selected as the standard performance of sensors. To improve the prediction accuracy of the semi-sparse pattern, which was preprocessed before putting images expressed from the semi-sparse pattern into the network training. There are differential and noise reduction treatments. The specific preprocessing methods are described as follows.

- Difference Process: Each set of deformation patterns has three images, the first is under no stress, the second is only under positive pressure, and the third is under a further shear force based on the second. As a result, the sensor causes thermal drift with a longer use time. To avoid this thermal drift, we get some differential graphs. The first graph of each group is taken minus the second graph, and the first graph is also taken minus the third graph, which is called the positive order difference. The opposite is the reverse order difference.
- Denoising Process: Denoising uses a thresholding operation. After trial and error, the threshold value of 30 both preserves most of the border information well and filters out a small part of the non-border noise. The threshold value is then set to 30, the pixel RGB value where the RGB value is added to less than 30 and the RGB value is set to 0.
- Original image: Not doing any processing is called the original image input.

As shown in Fig. 5, according to the above-preprocessing methods, five different preprocessing combinations were obtained: positive order difference without denoising (named A), reverse order difference without denoising (named B), positive order difference denoising (named C), and reverse order difference denoising (named D) and original image (named E).

After the above five preprocessing of the deformed images

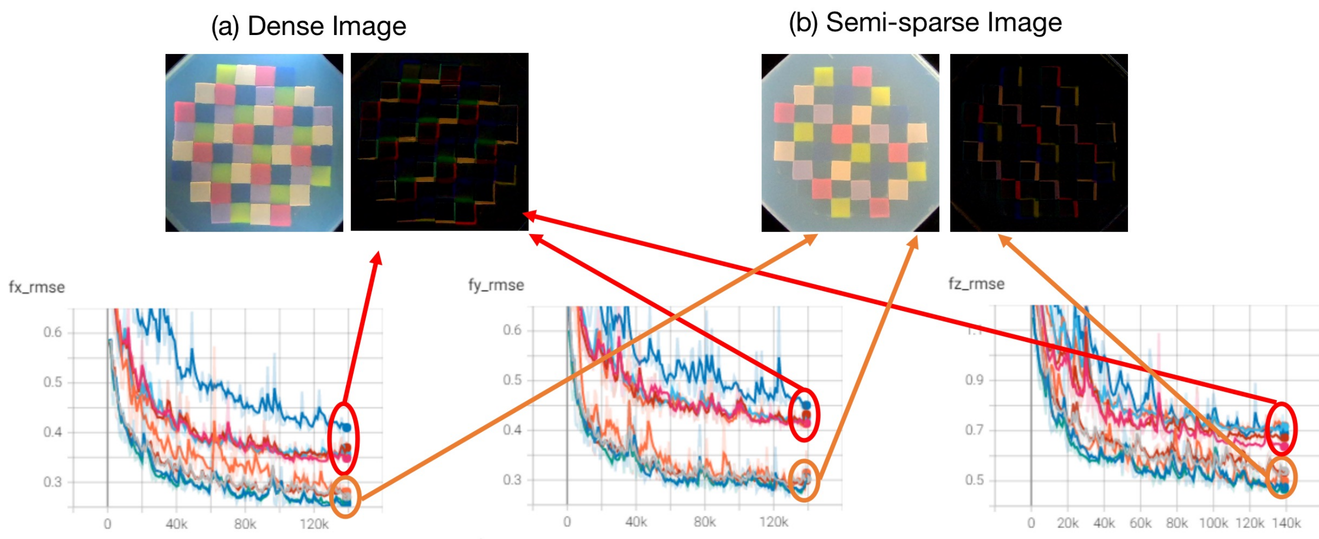
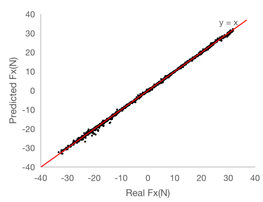
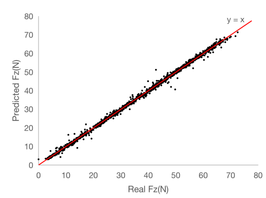


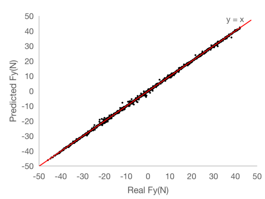
Fig. 7. The comparison of the semi-sparse and dense patterns in evaluating the five different preprocessing. The lines for dense pattern: orange line: preprocessing A, red line: preprocessing B, light blue line: preprocessing C, pick line: preprocessing D, blue line: preprocessing E



(a) F_x result



(b) F_y result



(c) F_z result

Fig. 8. The comparison of the force measured by the designed sensor and the ground truth in x, y and z direction. The red lines represent the ideal result, and the black dots represent measured result.

collected by the semi-sparse pattern, they were put into the Alexnet network [41] and evaluated with the validation set at each iteration. The graph represents the validation of F_x, F_y, F_z during the training process. The evaluation method used is RMSE, which for a good training network is as low as it should be possible. As seen from the Fig. 6, for the evaluation of F_x during the training process, B=D>A=C>E, for the evaluation of F_y during the training process, B=D>A=C>E, for the evaluation of F_z during the training process, B=D>E>A=C, it can be concluded that the preprocessing of the reverse order difference performs the best effect in evaluating the model for 3D force training, regardless of whether it is denoising or not. Otherwise, the positive difference performs very well on F_x and F_y, but not as well on F_z. Similarly, the dense pattern collected data for the above five preprocessing, pretreatment after the images into the same network training, get the training process of each cycle validation curve, as shown in the Fig. 7, under each preprocessing, semi-sparse pattern in F_x, F_y, F_z error are better than dense pattern, which also verifies the above conclusion: semi-sparse pattern contrast dense pattern has better resolution performance.

C. Sensor Evaluation

The best model of the semi-sparse difference non-denoising group is saved into the final trained network, and 2243 random data are collected to test the performance of the sensor network. The linear relationship between the predictive and actual values is studied. As shown in Fig. 8, black indicates the actual values, and red indicates the predicted values. It can see that the actual 3D force to predict the 3D force has a perfect linear relationship, indicating that the sensor in this paper has a superior force measurement capability. It further shows that the semi-sparse pattern can reach the resolution level of deep learning training.

The method mentioned enables accurate measurement of 3D forces, which outperforms conventional methods. Conventional resistive [7], capacitive [5] tactile sensors can achieve one-dimensional force (pull force or pressure) measurement, but this method is challenging to achieve in three-dimensional force detection. Because in the detection process, the planar shear force (F_x, F_y) and the vertical normal force (F_z) will cause the deformation of the sensor simultaneously, making the generated signals interfere with each other [25]. Through the method of structural innovation [28], there are still the problems of complex decoupling process, easy interference and low decoupling accuracy. It can be said that the tactile sensor of vision is superior to the tactile sensor based on the electrical and magnetic signal principle. The latter is vulnerable to electromagnetic signal interference and cannot accurately measure the force.

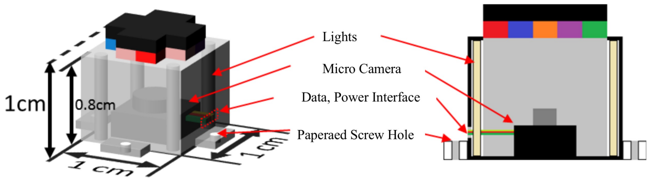
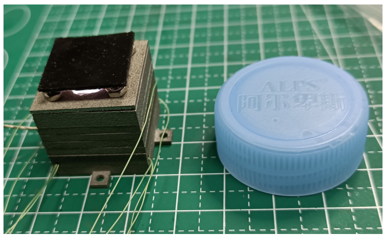
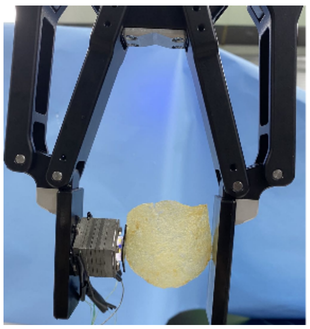


Fig. 9. Sensor Integration Scheme



(a) Manufactured sensor



(b) Chip-grasping Task

(c) Egg-grasping Task

Fig. 10. (a)The proposed sensor is easy to fabricate and the size of a bottle lid, enabling a wide range of applications. (b)(c)Robot gripper using tactile feedback from the proposed sensor to hold a chip/egg without squishing it.

D. Sensor Application

In terms of structural design, the design mentioned is streamlined. Silicone and camera, and light source are easy to integrate. At the same time, many other tactile sensors are very bulky and difficult to use. For example, in the optical fibre imaging proposed by [35], the optical fiber array consists of 121 single-core optical fibers. They are evenly interwoven and fixed to a metal plate measuring 40 cm in diameter at one end. The light source is connected to the other end of the input fiber, and the rear-end receiving camera is connected to the other end of the output fiber. However, their optical fiber is very bulky and not practical enough compared to designs in this paper, directly reflected in the silicone pattern design.

Therefore, by making customized miniature image acquisition equipment, component integration, reducing the size of structural parts and other means, the sensor size was controlled within $1cm^3$ (22cm $\times$22cm $\times$ 22cm), as shown in Fig. 9. At present, the detection frequency has reached 30Hz. Next, this paper conducts experiments to demonstrate the effectiveness of the proposed sensor in some application scenarios.

Force sensitivity: Water in a plastic bottle To visually illustrate the force sensitivity of our sensor, we do a pouring demo. By mounting the tactile sensor on a robotic gripper, which holds a plastic bottle, add water to the bottle several times without slipping off. During the process of pouring water, it can be seen that both the positive pressure and the friction force measurements of the sensor will increase, indicating the sensor's ability to distinguish forces as small as the weight of less than 20 mL ($<$0.2N) of water. (Supplementary Video S1)

Grasping fragile objects: Next, we show that the proposed sensor can be a sound tactile sensor for robotics applications such as grasping delicate objects such as chips and eggs, as shown in Fig. 10. Grasping squishy objects requires force feedback – too much force will squish the eggs and chips. This paper demonstrates that the built-in force induction (minimum of 30N) is insufficient for the task and that the proposed sensor does an excellent job of using force feedback to control grasping.

VI. CONCLUSIONS

The force resolution of 4 different pattern designs was compared using deep learning methods. The camera and light source imaging were used in the integrated system, keeping others consistent except for pattern design differences. By comparing the accuracy of deep learning decoupling forces, it can find that the force resolution of semi-sparse patterns is better than other pattern designs. Furthermore, using the grad-cam method, the focus of deep learning models on different pattern designs was obtained and found that the focus on semi-sparse pattern design tends to cover the whole image. Therefore, the design of semi-sparse patterns is more conducive to deep learning model learning. In this way, the connection between flexible probe pattern and deep learning image recognition mechanism is verified, which is conducive to the design of the flexible sensor. Subsequently, the semi-sparse pattern was selected as the study object, and its performance was analyzed. As a result, it found a good linear relationship between the predicted and actual values on the force measurement, indicating that the semi-sparse pattern met the resolution requirements of the deep learning training force. Finally, we integrated the sensor, making it small enough to install at the end of the mechanical gripper, and demonstrated the effectiveness of the proposed sensor through a series of experiments. For example, the experiment of adding water to the bottle was designed in terms of force sensitivity measurement. As a result, the force curve showed that both the positive pressure and the friction force were increased during the water-adding process. At the same time, sensors can also grasp fragile objects through force feedback, which cannot be achieved by built-in force induction. In the future, it is hoped to optimize the structure of the sensor to adapt it to more scenarios, and it can be integrated into electronic skin to accomplish more challenging tasks, such as adaptive grasping, human-computer interaction, Etc.

REFERENCES

[1] N. Akhtar and A. Mian, "Threat of adversarial attacks on deep learning in computer vision: A survey," *Ieee Access*, vol. 6, pp. 14 410–14 430, 2018.

[2] A. Billard and D. Kragic, "Trends and challenges in robot manipulation," *Science*, vol. 364, no. 6446, p. eaat8414, 2019.

[3] Q. Li, O. Kroemer, Z. Su, F. F. Veiga, M. Kaboli, and H. J. Ritter, "A review of tactile information: Perception and action through touch," *IEEE Transactions on Robotics*, vol. 36, no. 6, pp. 1619–1634, 2020.

[4] C. Bartolozzi, L. Natale, F. Nori, and G. Metta, "Robots with a sense of touch," *Nature materials*, vol. 15, no. 9, pp. 921–925, 2016.

[5] N. F. Lepora, U. Martinez-Hernandez, M. Evans, L. Natale, G. Metta, and T. J. Prescott, "Tactile superresolution and biomimetic hyperacuity," *IEEE Transactions on Robotics*, vol. 31, no. 3, pp. 605–618, 2015.

[6] C. M. Boutry, 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, no. 24, p. eaau6914, 2018.

[7] J. Park, M. Kim, Y. Lee, H. S. Lee, and H. Ko, "Fingertip skin–inspired microstructured ferroelectric skins discriminate static/dynamic pressure and temperature stimuli," *Science advances*, vol. 1, no. 9, p. e1500661, 2015.

[8] Z. Zou, C. Zhu, Y. Li, X. Lei, W. Zhang, and J. Xiao, "Rehealable, fully recyclable, and malleable electronic skin enabled by dynamic covalent thermoset nanocomposite," *Science advances*, vol. 4, no. 2, p. eaaq0508, 2018.

[9] H. Wang, G. De Boer, J. Kow, A. Alazmani, M. Ghajari, R. Hewson, and P. Culmer, "Design methodology for magnetic field-based soft tri-axis tactile sensors," *Sensors*, vol. 16, no. 9, p. 1356, 2016.

[10] T. P. Tomo, M. Regoli, A. Schmitz, L. Natale, H. Kristanto, S. Somlor, L. Jamone, G. Metta, and S. Sugano, "A new silicone structure for uskin—a soft, distributed, digital 3-axis skin sensor and its integration on the humanoid robot icub," *IEEE Robotics and Automation Letters*, vol. 3, no. 3, pp. 2584–2591, 2018.

[11] C. Ledermann, S. Wirges, D. Oertel, M. Mende, and H. Woern, "Tactile sensor on a magnetic basis using novel 3d hall sensor-first prototypes and results," in *2013 IEEE 17th International Conference on Intelligent Engineering Systems (INES)*. IEEE, 2013, pp. 55–60.

[12] H. Yousef, M. Boukallel, and K. Althoefer, "Tactile sensing for dexterous in-hand manipulation in robotics—a review," *Sensors and Actuators A: physical*, vol. 167, no. 2, pp. 171–187, 2011.

[13] N. F. Lepora, "Soft biomimetic optical tactile sensing with the tactip: a review," *IEEE Sensors Journal*, 2021.

[14] B. Ward-Cherrier, N. Pestell, L. Cramphorn, B. Winstone, M. E. Giannaccini, J. Rossiter, and N. F. Lepora, "The tactip family: Soft optical tactile sensors with 3d-printed biomimetic morphologies," *Soft robotics*, vol. 5, no. 2, pp. 216–227, 2018.

[15] A. Alspach, K. Hashimoto, N. Kuppuswamy, and R. Tedrake, "Soft-bubble: A highly compliant dense geometry tactile sensor for robot manipulation," in *2019 2nd IEEE International Conference on Soft Robotics (RoboSoft)*. IEEE, 2019, pp. 597–604.

[16] M. Lambeta, P.-W. Chou, S. Tian, B. Yang, B. Maloon, V. R. Most, D. Stroud, R. Santos, A. Byagowi, G. Kammerer *et al.*, "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, no. 3, pp. 3838–3845, 2020.

[17] D. Fehr, R. Sassenburg, J. Blunschi, A. Lay-Ekuakille, A. Massaro, M. Bonmarin, and F. Spano, "A capacitive color-changing electronic skin for touch sensing applications," in *2021 IEEE International Symposium on Medical Measurements and Applications (MeMeA)*. IEEE, 2021, pp. 1–6.

[18] A. C. Abad and A. Ranasinghe, "Visuotactile sensors with emphasis on gelsight sensor: A review," *IEEE Sensors Journal*, vol. 20, no. 14, pp. 7628–7638, 2020.

[19] R. Wang, C. Jiang, B. Yang, and J. Liu, "A novel vision-based tactile sensor using particle image velocimetry for multi-modal object detection and force sensing," in *2021 IEEE 34th International Conference on Micro Electro Mechanical Systems (MEMS)*, 2021.

[20] Z. Si and W. Yuan, "Taxim: An example-based simulation model for gelsight tactile sensors," 2021.

[21] G. Wang, "A perspective on deep imaging," *IEEE access*, vol. 4, pp. 8914–8924, 2016.

[22] K. Takahashi and J. Tan, "Deep visuo-tactile learning: Estimation of tactile properties from images," in *International Conference on Robotics and Automation*, 2019.

[23] W. Yuan, C. Zhu, A. Owens, M. A. Srinivasan, and E. H. Adelson, "Shape-independent hardness estimation using deep learning and a gelsight tactile sensor," in *IEEE*, 2017.

[24] P. Piacenza, S. Sherman, and M. Ciocarlie, "Data-driven super-resolution on a tactile dome," *IEEE Robotics and Automation Letters*, vol. 3, no. 3, pp. 1434–1441, 2018.

[25] C. Pang, G.-Y. Lee, T.-i. Kim, S. M. Kim, H. N. Kim, S.-H. Ahn, and K.-Y. Suh, "A flexible and highly sensitive strain-gauge sensor using reversible interlocking of nanofibres," *Nature materials*, vol. 11, no. 9, pp. 795–801, 2012.

[26] B. Nie, J. Geng, T. Yao, Y. Miao, Y. Zhang, X. Chen, and J. Liu, "Sensing arbitrary contact forces with a flexible porous dielectric elastomer," *Materials Horizons*, vol. 8, no. 3, pp. 962–971, 2021.

[27] Y. Zhao, W. Yang, Y. Shan, H. Bian, and M. Liu, "Design and calibration experiment of a novel rigid-flexible hybrid parallel three-dimensional force sensor with deformability," *IEEE Sensors Journal*, vol. 18, no. 20, pp. 8247–8255, 2018.

[28] Y. Song, F. Wang, and Z. Zhang, "Decoupling research of a novel three-dimensional force flexible tactile sensor based on an improved bp algorithm," *Micromachines*, vol. 9, no. 5, p. 236, 2018.

[29] Y. Zhu, S. Jiang, Y. Xiao, J. Yu, L. Sun, and W. Zhang, "A flexible three-dimensional force sensor based on pi piezoresistive film," *Journal of Materials Science: Materials in Electronics*, vol. 29, no. 23, pp. 19 830–19 839, 2018.

[30] S. Chen, C. Bai, C. Zhang, D. Geng, R. Liu, Y. Xie, and W. Zhou, "Flexible piezoresistive three-dimensional force sensor based on interlocked structures," *Sensors and Actuators A: Physical*, vol. 330, p. 112857, 2021.

[31] Y. Technology, "Tactile measurement with a gelsight sensor," 02 2015.

[32] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, "Bert: Pre-training of deep bidirectional transformers for language understanding," *arXiv preprint arXiv:1810.04805*, 2018.

[33] J. Oh, X. Guo, H. Lee, R. L. Lewis, and S. Singh, "Action-conditional video prediction using deep networks in atari games," *Advances in neural information processing systems*, vol. 28, 2015.

[34] T. N. Kipf and M. Welling, "Semi-supervised classification with graph convolutional networks," *arXiv preprint arXiv:1609.02907*, 2016.

[35] D. Baimukashev, Z. Kappassov, and H. A. Varol, "Shear, torsion and pressure tactile sensor via plastic optofiber guided imaging," *IEEE Robotics and Automation Letters*, vol. 5, no. 2, pp. 2618–2625, 2020.

[36] V. Kakani, X. Cui, M. Ma, and H. Kim, "Vision-based tactile sensor mechanism for the estimation of contact position and force distribution using deep learning," *Sensors*, vol. 21, no. 5, p. 1920, 2021.

[37] W. Yuan, R. Li, M. A. Srinivasan, and E. H. Adelson, "Measurement of shear and slip with a gelsight tactile sensor," in *2015 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2015, pp. 304–311.

[38] D. Heaven, "Deep trouble for deep learning," *Nature*, vol. 574, no. 7777, pp. 163–166, 2019.

[39] F.-L. Fan, J. Xiong, M. Li, and G. Wang, "On interpretability of artificial neural networks: A survey," *IEEE Transactions on Radiation and Plasma Medical Sciences*, vol. 5, no. 6, pp. 741–760, 2021.

[40] R. R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh, and D. Batra, "Grad-cam: Visual explanations from deep networks via gradient-based localization," in *Proceedings of the IEEE international conference on computer vision*, 2017, pp. 618–626.

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