Effect of pattern on the resolution of the visual-tactile sensor

Qingling Duan[1], Qi Zhang[2], Dong Luo[1] and Yongsheng Ou*[3]

***Abstract*— Thanks to high spatial resolution and multi-tactile mode perception, visual-tactile sensing technology has been widely used in various robot operations such as active perception, pose estimation and in-palm operation. However, the pattern as an essential part of the optic tactile sensor has rarely been studied. Therefore, this paper investigated the effect of the different patterns on the resolution of the visual-tactile sensors. The silicone sensors with different densities and sizes and cameras integrate different tactile sensors, collect the sensor stress data in different situations, and process the data with deep learning models. The performance of the prediction forces of different patterns was evaluated by the root mean square error (RMSE). The results show that the proposed four patterns can decouple the normal and shear forces, and the force resolution of the semi-sparse pattern is better than the other pattern designs. Furthermore, the grad-cam method is used to obtain the focus of the deep learning model decoupling forces, showing that the semi-sparse pattern tends to cover the whole image because it has better performance.**

I. INTRODUCTION

Thanks to the rapid development of computer vision, robots' visual perception and understanding ability of unstructured and natural scenes has rapidly improved [1]. However, when a robot has visual perception impairment, such as insufficient light supply and occlusion, in complex tactile contact task scenarios such as robot dexterous operation, good tactile feedback (such as contact force) can provide rich proprioception resulting in more reliable operation and control strategies. Therefore, designing soft force sensors like human skin is critical to the robot field, which can promote robot development.

The traditional soft tactile sensor generally detects the force signal through the change of capacitor [2,3], resistance [4,5] and other electrical signals caused by deformation under the action of external force. In the process of detection of the electrical signal, the plane shear force (F_x, F_y) and the normal force (F_z) in the vertical direction will simultaneously cause the deformation of the sensor, making the resulting signals interfere with [6] with each other and cannot decouple the three-dimensional force. However, with the development of vision algorithms, visual-tactile sensors have emerged, which have gradually become the hot spot direction of tactile sensors due to the advantages of high spatial resolution, low cost and rich tactile information.

[1]Qingling Duan and Dong Luo is with Faculty of Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences, Shenzhen 518055, Guangdong Province, PR China.

[2]Qi Zhang is with the Department of Shenzhen College of Advanced Technology, University of Chinese Academy of Sciences, Shenzhen, China

[3]Yongsheng Ou is with the Department of Guangdong Provincial Key Lab of Robotics and Intelligent System, Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences ys.ou@siat.ac.cn

The design of vision-tactile sensors comprises the colloidal contact layer, light source structure and camera imaging system. In MIT Gelsight's visual-tactile sensor [7], they introduced labeled points on the reflective film inside the soft elastomer to capture the displacement of labelled points under 3D forces and established the mapping relationship between the labelled point displacement and 3D forces through finite element analysis to realize 3D force detection in the soft environment. The TacTip series optic tactile sensor [8] was proposed by the University of Bristol Nathan et al. The TacTip sensor mimics the human fingertip touch-body receptor structure by embedding an array-distributed pin in the colloidal layer, thereby conducting the deformation information on the sensor surface using the camera system to observe the movement of the pin array. In addition, Sui et al. of Tsinghua University have developed the Tac3D tactile sensor[9], which contains an optical path system refracted by four light mirrors and can achieve a virtual binocular imaging effect through a monocular camera. Cui et al. of the Institute of Automation of the Chinese Academy of Sciences proposed the GelStereo visual-tactile sensing series [10] based on binocular vision, which can obtain binocular tactile images simultaneously and then recover the contact depth information through the stereo matching algorithm.

This paper designs a new sensor consisting of an elastomer, a camera placed at the bottom of the sensor and light sources. The elastomer contains a black, light-opaque layer at the top which is used to prevent interference from external light, a patterned layer in the middle, and a transparent layer at the bottom. When the sensor's surface is under stress, the elastomer will deform, causing the colour and intensity of the light emitted by the light source to change through the elastomer emission, and the image captured by the camera will change. This paper uses the patterned layer of different densities and sizes, and other factors remain consistent. Using the Alexnet network, the datasets collected from different pattern designs were processed, and RMSE evaluated the prediction performance of the forces to evaluate the effect of the patterns on the resolution of the sensor. The results show that the force resolution of the semi-sparse pattern is better than the other pattern designs.

II. 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

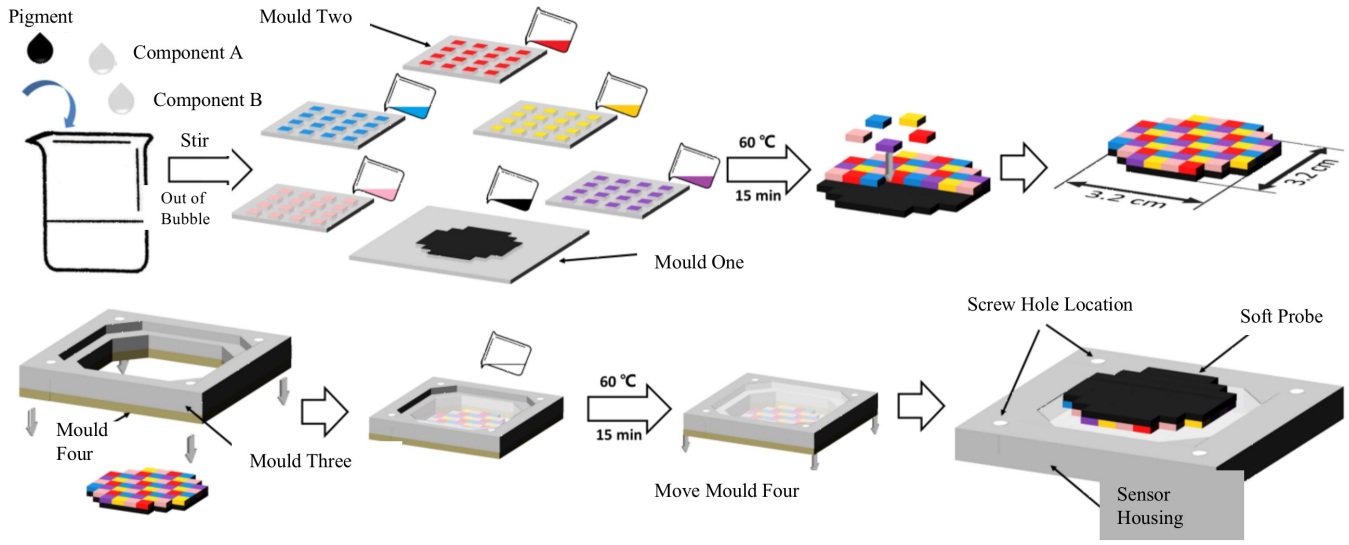


Fig. 1. The Soft Probe Preparation Process

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 pure Sorta-clear^{TM} 12 silicone as the transparent layer. The specific study method is shown in Fig. 1 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

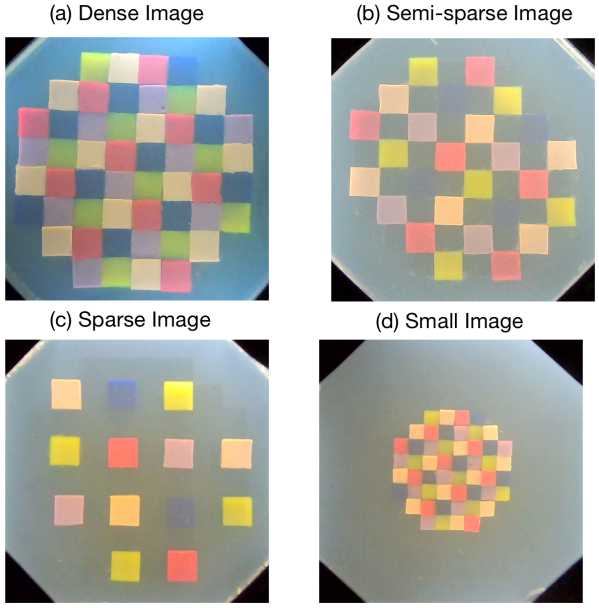


Fig. 2. Tactile Image Sensor Pattern Designs

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. 2(a), Fig. 2(b) and Fig. 2(c) differ in density, with Fig. 2(a) being the densest, Fig. 2(c) being the most sparse, and B being somewhere between them. Fig. 2(a) and Fig. 2(d) are used to contrast the effects of squares of different sizes. Therefore, based on the design in this paper, it only reflected the effect of patterns on the resolution of the

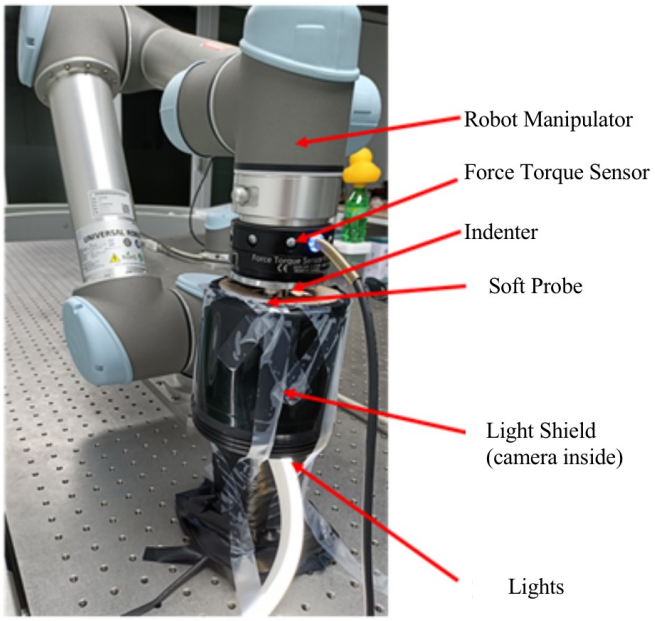


Fig. 3. Data Acquisition Platform

tactile sensor. The sparseness and miniaturization of patterns can help reduce the complexity of device production and accelerate the mass production of sensors.

III. 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 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 × 8 ×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

IV. Sensor Characterization

A. Effect of Pattern on the Resolution of Sensor

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 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 ± 0.03N, a F_y error between 0.43N ± 0.02N, a F_z error between and 0.68N ± 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, 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 = 20, F_y = 30, F_z = 60) and (F_x = -20, F_y = -10, F_z = 30). Due to the influence of the light source, it tends to cover the edge part (sunny position), which is obvious in the small pattern.

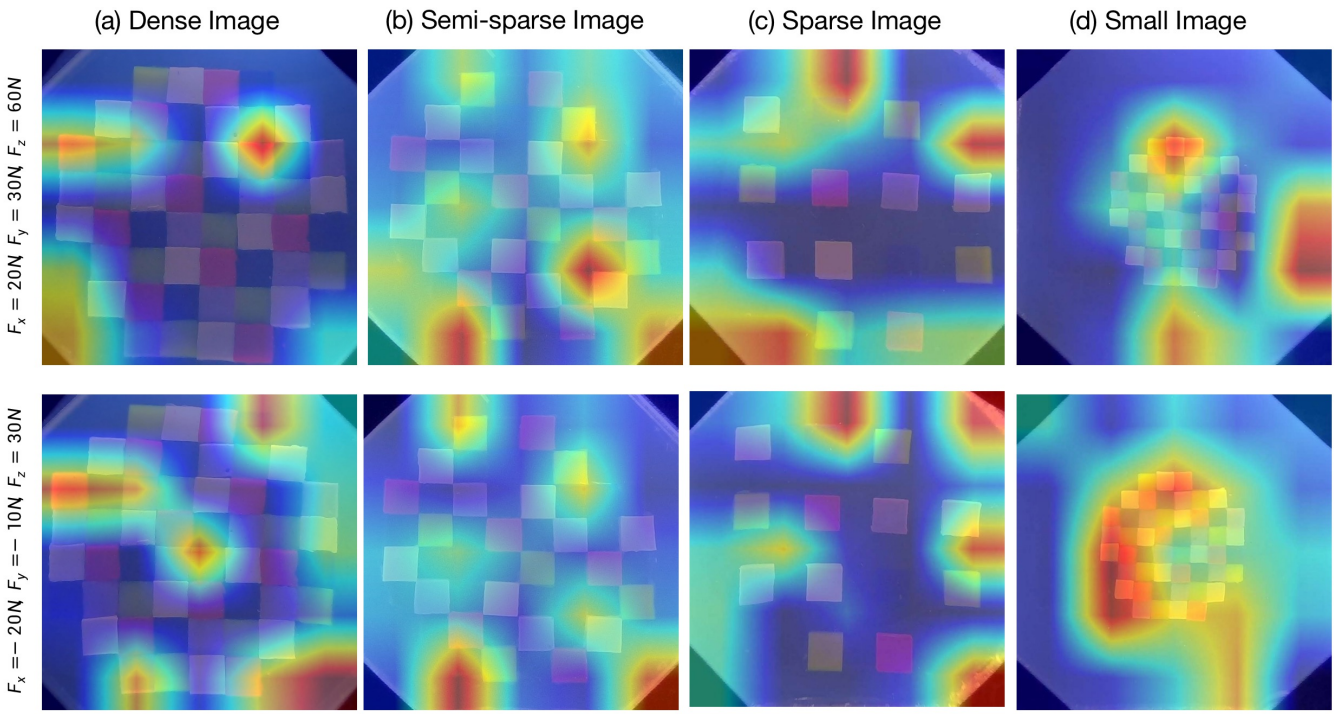


Fig. 4. The CAM Visualized using the Grad-CAM Method

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. Sensor Evaluation

Furthermore, 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. Differences are divided into positive and reverse order. The specific preprocessing methods are described as follows.

- Difference Process: Each set of deformation patterns has three images the first one is one with no force, the second one is only under positive pressure, and the third one is one with a shear force based on the second one. With the sensor, the independent image changes with time, so we will get the first graph of each group minus the second graph, and the first graph is also minus the third graph, which is called the positive order difference, and 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 collected by the semi-sparse pattern, they were put into the Alexnet network 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

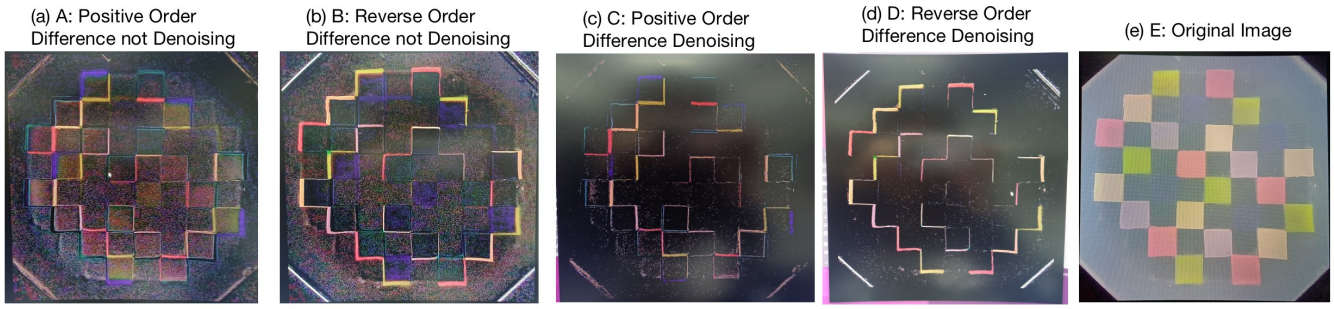
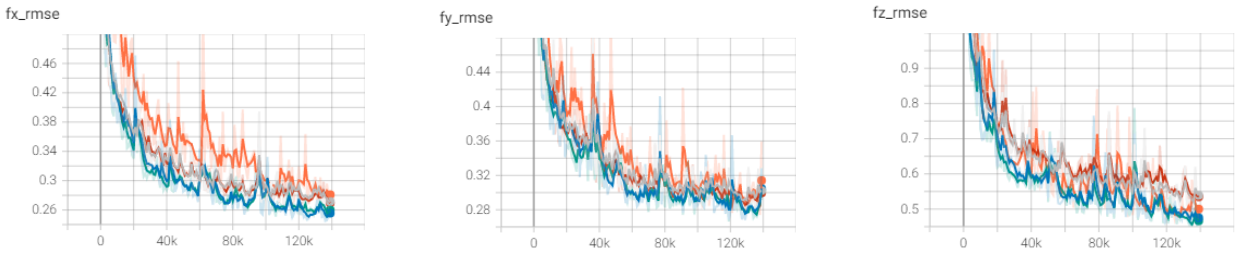


Fig. 5. Five Different Preprocessing Combinations



(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

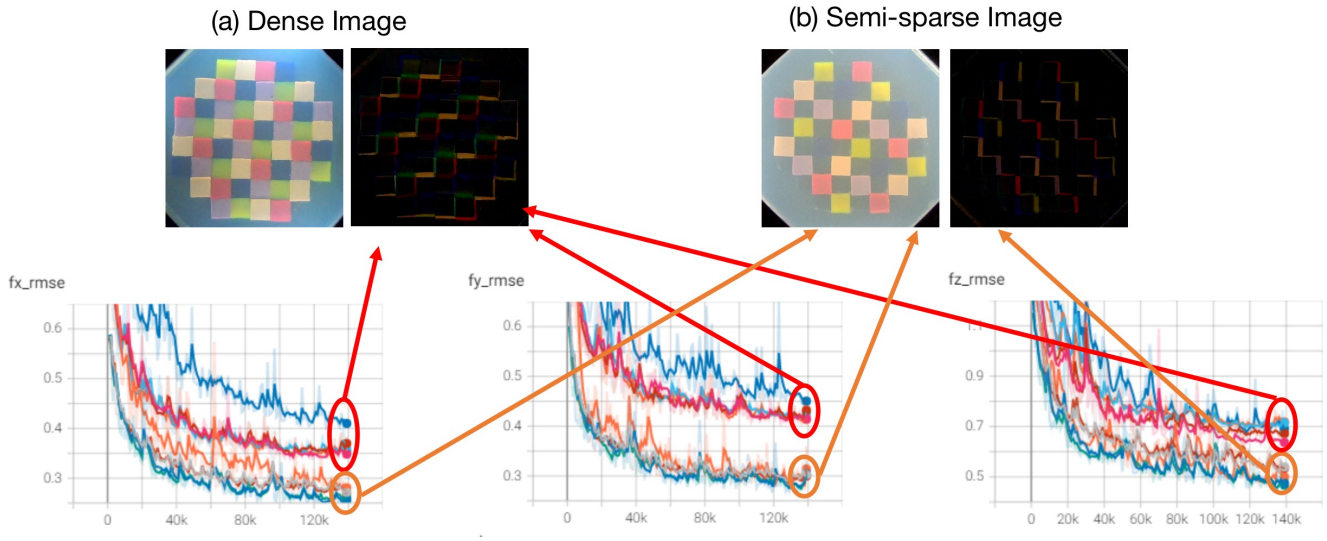
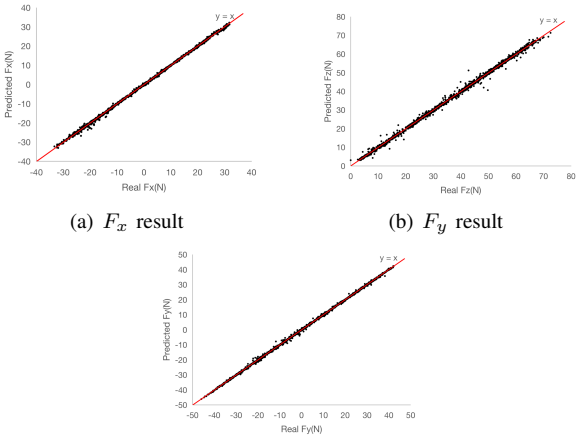


Fig. 7. The comparison of the semi-sparse and dense patterns in evaluating the five different preprocessing. The lines for dense pattern: orange line: preprocessing A, red line: preprocessing B, light blue line: preprocessing C, pick line: preprocessing D, blue line: preprocessing E

conclusion: semi-sparse pattern contrast dense pattern has better resolution performance.

Therefore, in this paper, 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, capacitive 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



(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.

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 the [6]. Through the method of structural innovation, 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 [11]. The latter is vulnerable to electromagnetic signal interference and cannot accurately measure the force.

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 [12], 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.

V. 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. Subsequently, the semi-sparse pattern was selected as the study object, and its performance was analyzed. Furthermore, 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. At the same time, the sensors are very conducive to integration. In the future, hoping to achieve the update and iterative development of soft 3D force sensors, hoping to be used in force control applications in flexible robot control.

ACKNOWLEDGMENT

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

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] Z. Shen, X. Zhu, C. Majidi, and G. Gu, "Cutaneous ionogel mechanoreceptors for soft machines, physiological sensing, and amputee prostheses," *Advanced Materials*, vol. 33, 09 2021.

[3] M. Fu, J. Zhang, Y. Jin, Y. Zhao, S. Huang, and C. Guo, "A highly sensitive, reliable, and high-temperature-resistant flexible pressure sensor based on ceramic nanofibers," *Advanced Science*, vol. 7, 09 2020.

[4] B. Ji, Q. Zhou, J. wu, G. Yibo, W. Wen, and B. Zhou, "Synergistic optimization toward the sensitivity and linearity of flexible pressure sensor via double conductive layer and porous microdome array," *ACS Applied Materials & Interfaces*, vol. XXXX, 06 2020.

[5] y. zhou, P. Zhan, M. Ren, G. Zheng, K. Dai, L. Mi, C. Liu, and C. Shen, "Significant stretchability enhancement of a crack-based strain sensor combined with high sensitivity and superior durability for motion monitoring," *ACS Applied Materials & Interfaces*, vol. 11, 01 2019.

[6] C. Pang, G.-Y. Lee, T.-i. Kim, S. 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, pp. 795–801, 07 2012.

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

[8] 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.

[9] L. Zhang, Y. Wang, and Y. Jiang, "Tac3d: A novel vision-based tactile sensor for measuring forces distribution and estimating friction coefficient distribution," *arXiv preprint arXiv:2202.06211*, 2022.

[10] S. Cui, R. Wang, J. Hu, J. Wei, S. Wang, and Z. Lou, "In-hand object localization using a novel high-resolution visuotactile sensor," *IEEE Transactions on Industrial Electronics*, vol. 69, no. 6, pp. 6015–6025, 2021.

[11] R. Bhirangi, T. Hellebrekers, C. Majidi, and A. Gupta, "Reskin: versatile, replaceable, lasting tactile skins," *arXiv preprint arXiv:2111.00071*, 2021.

[12] 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.