IEEE Robotics and Automation Letters (RA-L) paper, presented at ICRA 2024, Yokohama, Japan. Cite as RA-L paper.

# GelFinger: A Novel Visual-Tactile Sensor with Multi-Angle Tactile Image Stitching

Zhonglin Lin[1], *Member*, *IEEE*, Jiaquan Zhuang[1], Yufeng Li[1], Xianyu Wu[1], Shan Luo[2], Daniel Fernandes Gomes[2,3], Feng Huang[1], and Zheng Yang[1]

***Abstract*—Visual-tactile sensors that use a camera to capture the deformation of a soft gel layer have become popular in recent years. However, these sensors have a limited receptive field, which can hinder their ability to perceive tactile information effectively. In this paper, we propose a novel visual-tactile sensor named GelFinger that closely resembles the human finger and is well-suited for detecting various complex surfaces. The GelFinger sensor is equipped with an embedded miniature motor that allows for the adaptation of the camera pose and the scanning of a large contact area. During the detection process, the camera rotates to multiple angles to capture the tactile image of the contact area. To stitch together the tactile images obtained at different camera poses, we use an As-Projective-As-Possible image stitching algorithm to form a global view of the contact. We demonstrate the effectiveness of the GelFinger sensor in assessing large surfaces by using it to reconstruct curved crack outlines. Comparative experimental results show that the proposed sensor can effectively detect cracks and has the potential to assist humans in detecting defects on curved surfaces of infrastructure such as pipelines.**

***Index Terms*—Force and tactile sensing, perception for grasping and manipulation, object detection, segmentation and categorization**

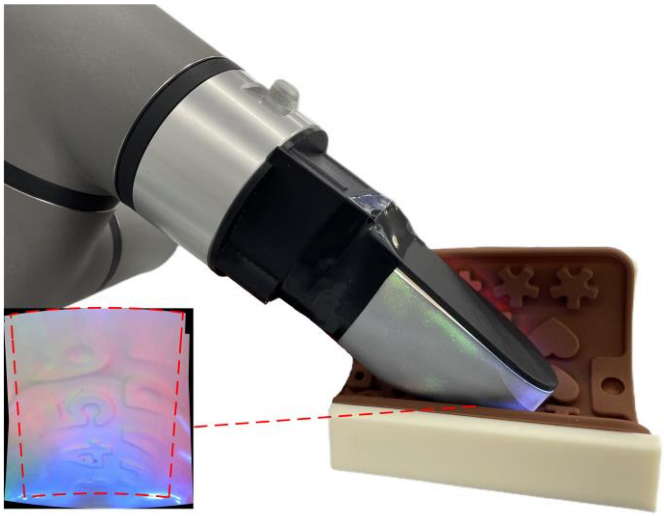

Fig. 1. GelFinger sensor, covered with a soft silicone pad, is used to assess a textured pipeline, with its output shown on the bottom left.

## I. Introduction

As robots continue to be utilized more frequently in human life and industrial applications, there is an increasing demand for higher performance from them. Tactile perception is a critical way for robots to perceive and interact with their environment, making it essential for task execution. Tactile information is also indispensable for research on robot dexterous manipulation, autonomous cognition, human-robot interaction, etc. At present, robots mainly use capacitive [1], photoelectric [2], or piezoresistive [3] tactile sensors to obtain tactile information. Compared to these sensors, a visual-tactile sensor visualizes the tactile information through a camera to obtain richer tactile information, such as texture, contact force, and other information.

In recent years, visual-tactile sensors have attracted increasing attention due to their low cost, high resolution, and strong anti-interference capability. Typically, these sensors consist of three components: an elastomer contact medium, LED lighting, and a camera. By capturing the elastomer deformation through the camera, high-resolution tactile images can be obtained, enabling the 3D reconstruction of the contact surface [4] and measurement of the contact force [5]. With the addition of deep learning, tactile sensors can perform various

Manuscript received: March 29, 2023; Revised: June 13, 2023; Accepted: July 20, 2023.

This paper was recommended for publication by Editor A. Banerjee upon evaluation of the Associate Editor and Reviewers' comments. This work was supported by the Funding of the Fuzhou University Research Start-up Funding (No. GXRC-20051), Funding of the 2020 Fujian Province Young and Middle-aged Teacher Education Research Project (Technology) (No. JAT200030), Funding of the Natural Science Foundation of Fujian Province of China (No. 2021J05113), and the EPSRC projects "ViTac: Visual-Tactile Synergy for Handling Flexible Materials" (No. EP/T033517/2). *(Corresponding author: Xianyu Wu.)*

[1]Zhonglin Lin, Jiaquan Zhuang, Yufeng Li, Xianyu Wu, Feng Huang, and Zheng Yang are with the School of Mechanical Engineering and Automation, Fuzhou University, Fuzhou, Fujian, 350116 China (e-mail: linzhonglin@gmail.com; zhuangjq0105@163.com; liyufengfzu@163.com; xianyuwu@139.com; huangf@fzu.edu.cn; yz97527@163.com).

[2]Shan Luo and Daniel Fernandes Gomes are with the Department of Engineering, King's College London, WC2R 2LS London, U.K. (e-mail: {shan.luo, daniel.gomes}@kcl.ac.uk).

[3]Daniel Fernandes Gomes is also with the smARTLab, Department of Computer Science, University of Liverpool, L69 3BX Liverpool, U.K. (e-mail: danfergo@liverpool.ac.uk).

Digital Object Identifier (DOI): see top of this page.

IEEE Robotics and Automation Letters (RA-L) paper, presented at ICRA 2024, Yokohama, Japan. Cite as RA-L paper.

IEEE Robotics and Automation Letters (RA-L) paper, presented at ICRA 2024, Yokohama, Japan. Cite as RA-L paper.

applications such as detecting contact surface texture [6] and multi-sensor information fusion for fine operation tasks [7]. Despite the development of many visual-tactile sensors such as GelForce [8], GelSight [9], and DIGIT [10], most sensors have a small contact surface area and are mostly flat, limiting their detection capabilities to regular flat surfaces. When faced with a large detection surface, these sensors may struggle to provide effective detection. To address this issue, Cao et al. [11] proposed a sensor called TouchRoller, which features a rollable cylindrical design for fast, continuous detection of large areas. However, due to its structural design, TouchRoller is difficult to detect on irregular or undulating surfaces.

The goal of our work is to develop a visual-tactile sensor capable of efficient large-area detection on both planar and curved surfaces. To achieve this, in this paper, we present a new sensor design named GelFinger, as shown in Fig. 1, which can acquire more tactile information with a single touch. Our visual-tactile sensor is versatile and can be applied to various application scenarios such as robot dexterous manipulation and pipeline inspection. Our work offers three main contributions to this field:

- We introduce the GelFinger sensor, a novel visual-tactile sensor with a curved contact surface that resembles the shape of the human finger. The silicone of the sensor covers an area of approximately 74.6 $cm^2$. After the sensor makes contact with an object, the internal camera sequentially captures five photos. Then, five photos are stitched together, and the stitched picture can cover an area of 15 $cm^2$. The design enables the sensor to meet the detection requirements of curved surfaces.
- We have designed a miniature transmission module within the GelFinger sensor, which enables the camera to rotate and acquire tactile images of the contact area from various angles. This approach overcomes the challenge of poor imaging quality when detecting curved surfaces using tactile sensors.
- Using the As-Projective-As-Possible (APAP) image stitching method, we can obtain a large area of detection image information with just one contact. The stitched image is then used to detect cracks in pipelines using the DeepLabv3+ algorithm with ResNet as the backbone network. Our experimental results show that the proposed sensor has achieved an impressive MIoU of 90.9% in detecting pipeline cracks.

The structure of the paper is outlined as follows: Section II provides a comprehensive review of related works on visual-tactile sensor designs. Section III details the design and manufacturing process of the GelFinger sensor. In Section IV, we present the image stitching method used to combine the data collected from the sensor and showcase the stitching results from the GelFinger sensor. Section V focuses on demonstrating the detection performance of the sensor, including its ability to efficiently detect large areas. In Section VI, we introduce our pipeline crack detection algorithm and provide details on the tactile dataset and model training. With the results of our crack detection experiments. Finally, Section VII offers concluding remarks, highlights the key contributions of our work and outlines optional directions for future research.

## II. Related Work

### A. Visual-tactile Sensors

GelSight is a notable research branch in the field of visual-tactile sensors. The retrographic sensor, first introduced by Johnson and Adelson [12] in 2009, served as the basis for the development of GelSight. This novel sensor is capable of converting information about surface shape and pressure into images. Johnson et al. later created an improved version of the retrographic sensor in 2011 [13], which was designed to be portable and handheld. In 2014, Li et al. [14] proposed a cube design for the GelSight sensor. Further improvements to GelSight were introduced in 2017 and a hexagonal prism-shaped sensor was proposed [9].

There is a diverse range of visual-tactile sensors that have been developed building on the foundation of GelSight but with cosmetic and structural modifications to cater to specific functional requirements. For instance, Gomes et al. [15] [16] developed a finger-type visual-tactile sensor known as GelTip. The sensor is constructed by covering the surface of a transparent tube with a metal coating of silicone, allowing it to sense contact anywhere on its surface. Similarly, Do and Kennedy [17] proposed a hemispherical tactile sensor called DenseTact, which is capable of estimating the sensor surface deformation in real time using convolutional neural networks. Remero et al. [18] developed a visual-tactile sensor with a curved contact surface that uses directional illumination in the form of a light tube and can capture high-resolution contact signals from contact surfaces. Sun et al. [19] propose a thumb-sized tactile sensor named Insight, which is capable of sensing the full range of force through vision and machine learning. However, all these sensors suffer from a limited receptive field, making them ineffective in assessing large surfaces.

### B. Crack Detection

Cracks are a pervasive issue in infrastructures such as pipelines and roads, posing serious safety risks. Early detection of cracks is crucial for their maintenance and safety. Traditional manual methods of crack detection are associated with high time consumption, cost, and subjectivity. Visual detection technology has emerged as the most common method of crack detection and recent advancements in deep neural networks have enabled pixel-level detection. In [20], end-to-end multi-scale full-convolutional and deconvolutional neural are developed for pixel-level detection. By learning crack features in the complex fine-grained background of asphalt pavements, semantic segmentation of richer multi-scale crack feature information is realized. In [21], an improved DeepLabv3+ network is proposed for denser pixel sampling, achieving higher accuracy in crack segmentation. However, vision-based crack detection methods are susceptible to external influences and can be challenging to use in special environments.

In this study, we propose a new approach to pipeline crack

IEEE Robotics and Automation Letters (RA-L) paper, presented at ICRA 2024, Yokohama, Japan. Cite as RA-L paper.

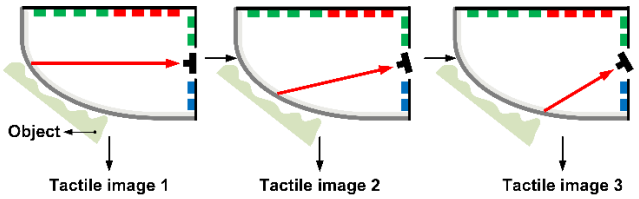


Fig. 2. Workflow of the GelFinger sensor.

detection that is based on the GelFinger sensor. By using tactile images of pipelines captured by the GelFinger sensor, this approach has shown promising results and proves that the GelFinger sensor can be a valuable tool for detecting cracks in pipelines.

## III. Design And Manufacture

### A. Working Principle of GelFinger Sensor

Visual-tactile sensors commonly consist of several key components including a camera, LED lights, a transparent acrylic support plate, a transparent silicone elastomer, and a metal reflective coating. The LED lights inside the sensor provide illumination, while transparent silicone elastomer and reflective metal coating act as the contact medium for tactile sensing. Additionally, a transparent acrylic plate serves as a support for the sensor. When the elastomer comes into contact with an object, it deforms according to the object's surface geometry, capturing tactile information that can be visualized through the camera.

The workflow of the GelFinger sensor is shown in Fig. 2. When the sensor comes in contact with an object, the imaging module is positioned to capture the initial tactile information, resulting in the first tactile image. The imaging module then rotates at a certain angle to capture the second tactile image and continues to rotate to obtain additional tactile images from multiple angles and positions. Once a set of tactile images has been captured, they are stitched together using an image stitching algorithm to create a complete tactile image to enable the detection of larger areas of the object being sensed.

### B. Sensor Structure Design

The structural design of the GelFinger sensor proposed is shown in Fig. 3. The sensor's contact surface is predominantly the fingertip curved contact surface, which comprises a metal reflective coating, a transparent silicone layer, and a transparent support layer from the outermost to innermost layer. The sensor integrates various components, including an imaging module, LED illumination module, micro motor transmission module, and an external control circuit module. The LED illumination module consists of two U-shaped PCBs, located in the inner plane of the sensor and the side plane of the camera, respectively, and contains red, green, and blue LEDs to ensure adequate illumination. The camera is positioned at the center of the LED light module and is driven by a micro motor, enabling the imaging module to capture images from any angle. Finally, the tail of the sensor is designed with a manipulator connector to enable installation on a manipulator for detection.

GelFinger's fingertip features a uniquely designed curved contact surface that allows it to adapt to a wide variety of

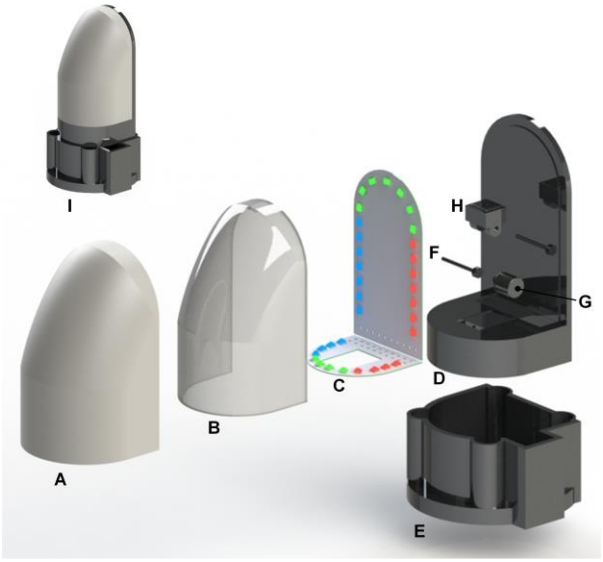


Fig. 3. The GelFinger is comprised of several key components: A) A specialized fingertip with a coating layer made from reflective metals and a transparent silicone layer; B) A transparent support layer that provides structural support to the fingertip and helps to distribute pressure evenly across its surface; C) An RGB LED lighting system that illuminates the fingertip; D) A sensor housing; E) A connector for a robotic arm that allows GelFinger to be seamlessly integrated into larger robotic systems for enhanced functionality; F) A drive shaft and bearing that allow for precise and smooth movements of the imaging module; G) A micro motor; H) A camera and mounting system. I) A 3D model of GelFinger.

detection environments, including uneven surfaces and the interiors of pipelines and other complex shapes. Unlike many existing sensors with planar contact surfaces that are limited to detecting only planar contact surfaces or objects with regular geometric shapes, GelFinger can detect irregular shapes and surfaces with greater accuracy and flexibility.

### C. Optical System Design

The GelFinger sensor operates on the principle of receiving tactile information through an optical system that is crucial to the sensor's performance. The optical system consists of a camera module, an RGB LED illumination module, and a micro motor transmission module. The sensor's metal reflective coating and housing isolate external light sources, ensuring that the sensor is illuminated solely by the LEDs, which provide a constant level of brightness. The internal imaging system of the GelFinger sensor is driven by a micro motor transmission module that rotates the imaging module to multiple angles. The LED light source within the sensor refracts and reflects when it hits the metal reflective coating, producing tactile information that is captured by the internal imaging module.

During the development of the GelFinger sensor, two design options were considered for the LED lighting module. The first option involves a micro-motor-driven LED lighting module that rotates together with the camera to direct the LED light to the contact point for illumination. The second option, detailed in Section III-B, was tested against both options in an experiment using a 3D-printed sphere as the detection object and the same camera for both options. The LED lighting module design and the experimental results, shown in Fig. 4,

IEEE Robotics and Automation Letters (RA-L) paper, presented at ICRA 2024, Yokohama, Japan. Cite as RA-L paper.

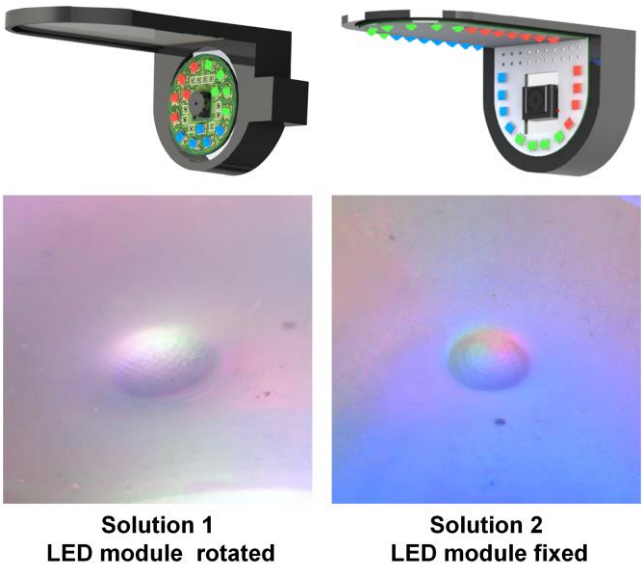


Fig. 4. Experimental results of two LED lighting module design solutions. As shown in the figure, the tactile image on the right provides a higher image contrast and greater gradients compared to the left one.

indicated that the second solution was superior to the first one. The second option results in better light distribution. Upon analysis, we determine that the LED light of the first option shines directly onto the contact point, causing a lower contrast image with less surface gradient information. In addition, this solution results in a change in the angle of incidence of the RGB LED lights due to the rotation of the LED lighting modules. In contrast, the LED of the second option shines from the side, providing higher image contrast and more gradient information. Based on these findings, we adopted the LED light fixation scheme of the second option for the final design of the GelFinger sensor.

### D. GelFinger Sensor Fabrication

In the development of GelFinger, fabricating an elastomer layer of fingertip-type surfaces with good transparency, resilience, hardness, and viscosity presents a significant challenge. We focus on finding a hybrid solution that could meet these requirements. Through experimentation, we found that using a blend of Smooth-On's Solaris parts A and B, and Slacker in 1:1.5:1 produced an elastomer with optimal hardness, transparency, viscosity, and resilience. To create the specially shaped silicone elastomer with a fingertip curved surface, we developed a molding process. The mold is composed of two layers, and once the silicone is poured and solidified, the upper and lower parts of the mold are separated to obtain the desired fingertip-type curved silicone shape.

Achieving a good fit between the metal reflective coating and the transparent silicone elastomer, while also ensuring the coating is opaque, is crucial in the fabrication. To prepare the coating, we mixed Smooth-On's Psycho Paint A and B, aluminum powder, and Smooth-On's Cast Magic Silver Bullet in a ratio of 1:1:0.3:0.2, and dissolved the mixture using Smooth-On's NOVOCS Matte. The aluminum powder with a diameter of $4\times10^{-6}$ m from a company named Zuxing. Finally, we sprayed the coating with a spray gun. The transparent

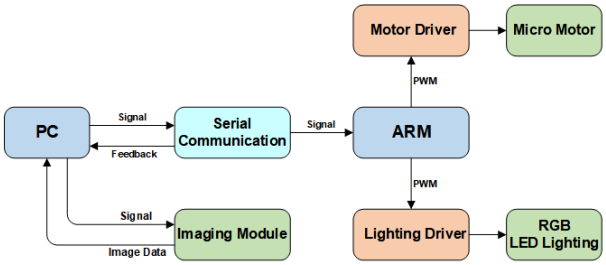


Fig. 5. Control system of the GelFinger.

support layer of the sensor was 3D printed, using VeroClear materials from Stratasys. After finishing printing, it needs to be sanded and polished, and varnish is sprayed on its surface. The housing part of the sensor was also 3D printed, and the completed preparation and assembly of each part can be seen in Fig. 3.

### E. Control System

The imaging module of the sensor is the OmniVision OV5640 image sensor, which captures images through the USB camera controller AU3841 and transmits the images to a PC via USB. The communication between the PC and the STM32F407ZET6 ARM processor is facilitated by the serial communication module (CH340), which sends control signals and receives feedback signals from the ARM processor. The ARM processor processes the output of Pulse Width Modulation (PWM) signals and controls the motor drive module (L298N) and the LED driver module (ULN2003) according to the received instructions. The micro stepper motors and the LED lighting circuits are driven independently. A software package has been developed to control the sensor, and the components of the control system are shown in Fig. 5.

The motor control is a crucial aspect of the sensor's operation. During the acquisition process, the robotic arm controls the sensor to make contact with the object being measured. Since there is no external camera assistance, the contact point is predetermined or manually determined. As shown in Fig. 2, the micro stepper motor drives the camera to rotate downward for 5° and acquires the first picture, which takes about 3-4 s from the start of the motor rotation to the acquisition of the first picture. The above steps are repeated to capture 5 images for a total of about 15-20 s. The GelFinger is mounted on the robot arm to make contact with the object, and the sensor performs a small angle rotation with the rotation speed matching the internal camera. The five images are then stitched together using an image stitching algorithm detailed in Section IV.

## IV. Image Stitching

### A. Image Stitching Algorithms

The sensor employs advanced image stitching technology to stitch tactile images captured from multiple angles, allowing it to detect a single, large area with high-quality image quality. Two different stitching methods were employed in the experiment: the image stitching method based on the camera calibration, and the APAP image stitching algorithm.

The stitching method based on camera calibration uses the camera calibration information to calculate the image

IEEE Robotics and Automation Letters (RA-L) paper, presented at ICRA 2024, Yokohama, Japan. Cite as RA-L paper.

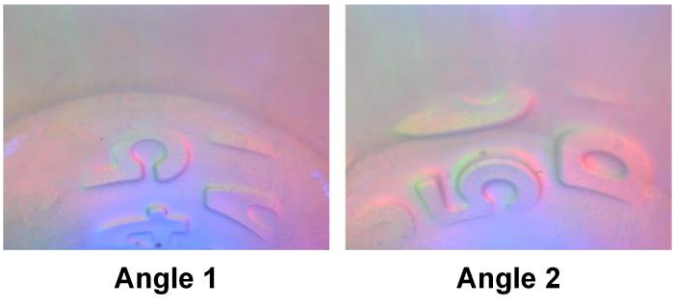


Fig. 6. Significant parallax in the tactile images obtained from different angles.

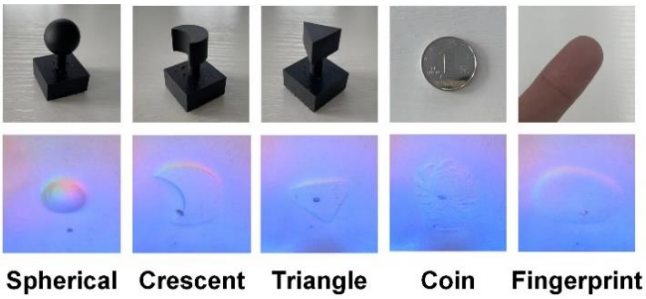


Fig. 7. Top: Objects used to evaluate the performance of GelFinger: from left to right, 3D printed spherical, crescent, and triangle shapes, coin, and a human fingertip; Bottom: Tactile image collected from each object by the GelFinger sensor.

TABLE I
SENSOR SHAPE AND SILICONE AREA

| Sensor | Shape of sensor | Silicone area |
|---|---|---|
| GelFinger | Fingertip shape | 74.62 $cm^2$ |
| GelSight | Hexagonal shape | 6.88 $cm^2$ |
| DIGIT | Prismatic shape | 3.87 $cm^2$ |
| GelTip | Finger shape | 42.41 $cm^2$ |

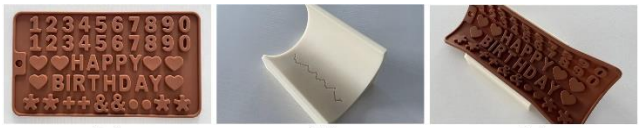


Fig. 8. The objects used for calculating the detection area.

alignment parameters and determine the projection transformation relationship between images. In the subsequent formal stitching, only the projection map and fusion based on the obtained transformation relationships are required to obtain images with large fields of view.

The APAP image stitching algorithm [22], on the other hand, is a local transformation method that is particularly effective when dealing with significant parallax between input images. It works by dividing the image into $C_1 \times C_2$ grids and calculating the corresponding local homography matrix using the Moving Direct Linear Transformation (MDLT) method. For the local homography matrix $\hat{\boldsymbol{h}}^k$ to be estimated, the calculation formula is as (1), and the final required local transformation matrix $\hat{\boldsymbol{h}}^k$ can be obtained by solving the eigenvector corresponding to the minimum singular value:

$$\begin{aligned} \hat{\boldsymbol{h}}^k &= \arg\min_{\boldsymbol{h}} \sum_{i=1}^{n} \left\| w_i^k \boldsymbol{a}_i \boldsymbol{h} \right\|^2 = \arg\min_{\boldsymbol{h}} \left\| \boldsymbol{W}^k \boldsymbol{A} \boldsymbol{h} \right\|^2 \\ \boldsymbol{W}^k &= diag([w_1^k \quad w_1^k \quad w_2^k \quad w_1^k L \quad w_n^k \quad w_n^k]) \end{aligned} \tag{1}$$

where $\boldsymbol{a}_i$ is the linear parameter matrix $\|\boldsymbol{h}\|$ corresponding to the converted matching feature point variable relationship, and it is specified as 1. $n$ is the number of matching point pairs, and $w_i^k$ is the corresponding weight of feature points $(x_i, y_i)$. The solution for $w_i^k$ is as follows:

$$\begin{aligned} w_i^k &= \max(\exp(-\frac{((x^k - x_i)^2 + (y^k - y^i)^2)}{\sigma^2}), \gamma) \\ &, \gamma \in (0,1), k \in [1, C_1 \times C_2] \end{aligned} \tag{2}$$

where $\sigma$ is the Gaussian scale factor, $\gamma$ is the threshold value set by ignoring feature points that are too far from the center of the grid, and $x^k$ and $y^k$ are coordinate values corresponding to the center of the $k$-th grid.

The image is then transformed and fused using the local homography matrix to complete the stitching process. Compared with the traditional stitching methods, the APAP algorithm is better suitable for dealing with local deformation and the presence of parallax and offers superior registration accuracy and adaptability.

### B. Algorithms Comparison and Results

To compare the effectiveness, speed, and practicality of the two image stitching methods, we first used the camera calibration based image stitching method which uses pre-calculated projection parameters for stitching, and thus stitching is fast and also achieves relatively good results. During the stitching process, we found that the camera rotating around the motor drive axis, which resulted in a distortion of the contact surface captured by the camera. The resulting contact points on the curved surface were different and closer to the camera, leading to significant parallax in the acquired tactile images. As shown in Fig. 6, the tactile images captured from different angles demonstrate this large parallax issue. To address this problem, we tried the APAP stitching method. The use of the APAP algorithm enables a better stitching effect, which is demonstrated in the following tests. Although this method is able to achieve better stitching results and solve the parallax problem, it is relatively slow due to the partitioning of the input image into grids and the calculation of the local homography matrix for each grid separately.

Compared with the camera calibration based image stitching algorithm, APAP can be more flexible in stitching, while the camera calibration based image stitching algorithm needs to be recalibrated for different stitching scenarios, and considering the need to perform more complex detection tasks with sensors in the future, we believe that the APAP image stitching algorithm has superior practical application performance, and therefore choose it as our stitching algorithm.

## V. SENSOR PERFORMANCE

### A. Experiment Setup

To evaluate the performance of the GelFinger sensor, we compared it to three other sensors in this study. Firstly, we produced a sensor with the same size and shape as the GelFinger sensor, but with a 160° field of view in the imaging module, and the imaging module was fixed and could not be rotated. This sensor may result in stretching and distortion of the detection images. Secondly, we built a sensor with the same size and shape as the GelFinger sensor, but with a 60° field of

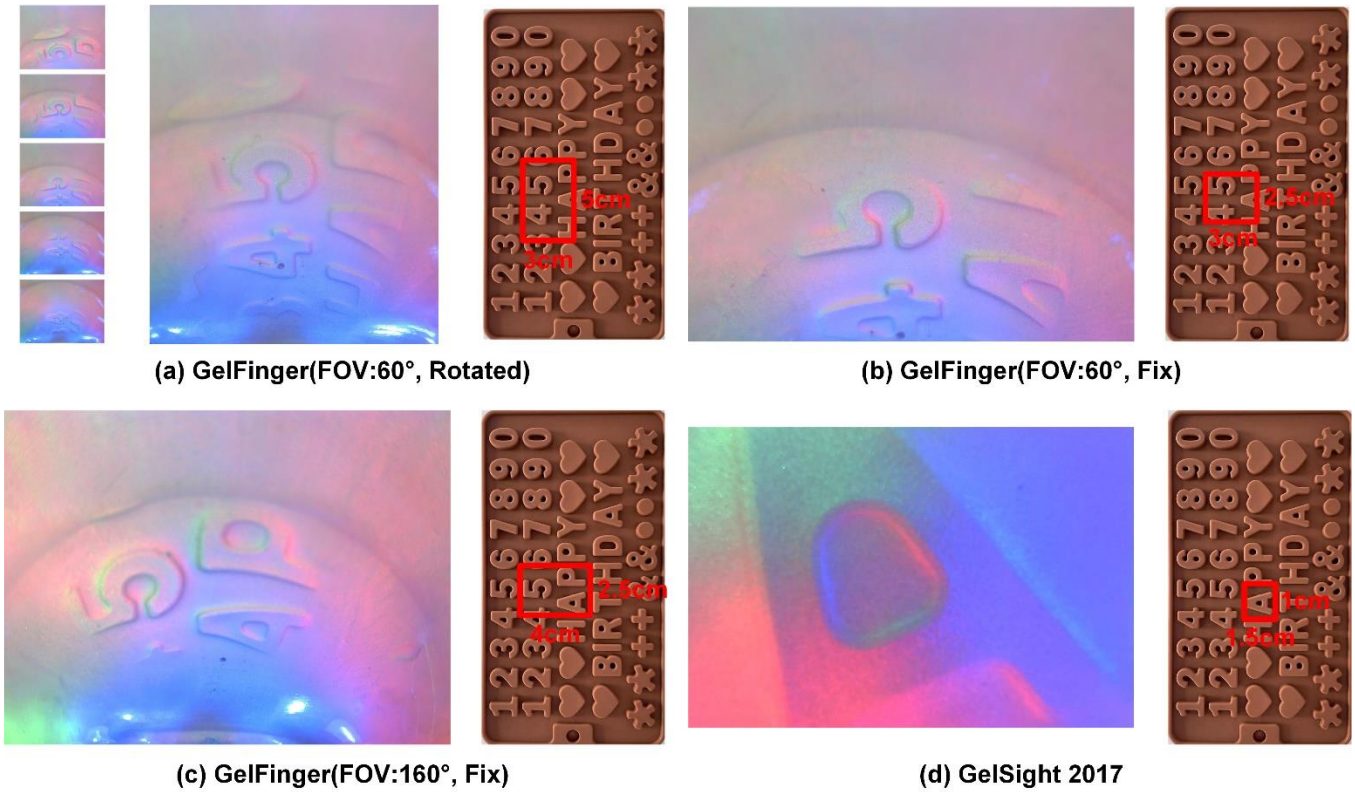


Fig. 9. Test results of the four sensors to detect the letter silicone model, in which the detection results of the sensors are shown on the left and the corresponding single detection areas are shown on the right.

TABLE II
DETECTION AREA

| Sensor | Camera | Detection area |
|---|---|---|
| GelFinger | FOV: 60°, Rotated | 15.0 $cm^2$ |
| GelFinger | FOV: 60°, Fixed | 7.5 $cm^2$ |
| GelFinger | FOV: 160°, Fixed | 10.0 $cm^2$ |
| GelSight | FOV: 60°, Fixed | 1.5 $cm^2$ |

view in the imaging module, and the imaging module was fixed and could not be rotated. This sensor is the typical configuration used in most optical tactile sensors with curved or flat detection surfaces. Thirdly, we produced the GelSight 2017 sensor for comparison.

To test the detection performance of the four sensors, we prepared various 3D printed objects as well as a pipeline model, to evaluate the performance of different sensors for a single detection area. In addition, we used coins and fingerprints to assess the detection performance of the sensors on fine textures.

### B. Shape and Silicone Area

We analyze and compare the shape and silicone area of the sensors, with the main analysis presented in Table I. In addition to GelFinger and GelSight sensors, we included DIGIT and GelTip sensors as a comparison based on the information from existing literature.

The comparison focused on four sensor types, including two with flat silicone detection surfaces and two with curved surfaces. The sensors with flat surfaces have smaller silicone detection areas and overall sensor sizes, while those with curved surfaces have larger overall sizes and a wider range of surface types compared to the flat sensors. Our GelFinger sensor includes a motor drive module, resulting in a larger overall size and the largest silicone contact area among the four sensor types. While this size may limit its use in small space environments, under appropriate operating conditions it can achieve fast and efficient detection of flat and curved surfaces.

### C. General Detection Performance

The GelFinger sensor was mounted on the Universal Robots UR3 robotic arm to conduct experiments and evaluate its ability to detect objects with conventional geometry. We prepared various models including spherical, crescent, triangle, coin, and fingertip shapes as shown in Fig. 7. The sensor was able to accurately represent the surface geometry of each object, and was also able to capture texture information from fine surfaces such as fingerprints and coins.

### D. Detection Area

To verify the GelFinger sensor's ability to detect large areas in a single detection, we conducted experiments with the four sensors described in Section V-A. The objects for calculating the detection area are shown in Fig. 8, including a letter silicone model, a pipeline model, and a combination of the two models. During the experiment, the sensor was used to obtain tactile images containing letters shown in Fig. 8(c), The tactile image acquisition method is shown in Fig. 1. The detection area of the sensor for a single detection was calculated by comparing the tactile image size with the real model size.

First, we use the GelFinger sensor to obtain tactile images taken by the camera at five different angles. These images are stitched together and the obtained results are shown in Fig. 9(a). For the silicone letter model, the area of a single detection was

IEEE Robotics and Automation Letters (RA-L) paper, presented at ICRA 2024, Yokohama, Japan. Cite as RA-L paper.

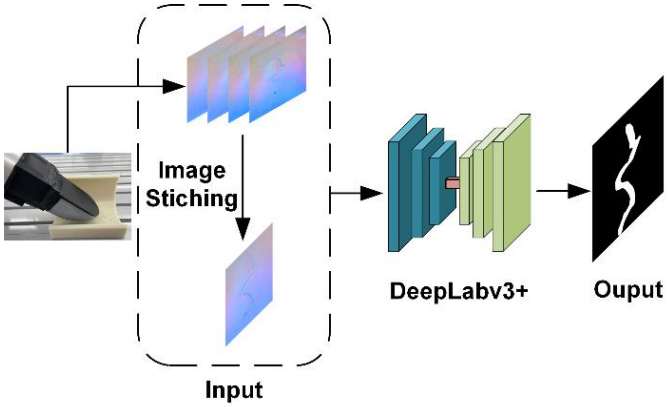


Fig. 10. Using the DeepLabv3+ to identify and segment cracks in tactile images.

TABLE III
DIFFERENT BACKBONE NETWORK PERFORMANCE

| Backbone | MIoU |
|---|---|
| ResNet | 90.9% |
| MobileNet | 89.6% |

calculated to be 15 $cm^2$. The detection areas of the remaining three sensors were calculated using the same method. Since GelSight sensors are less effective for detecting curved surfaces, GelSight was used for detecting flat surfaces. The experimental results are shown in Fig. 9(b-d) with quantitative results shown in Table II.

Our experimental results prove that the GelFinger sensor compared to visual-tactile sensors using ordinary cameras and wide-angle cameras, can achieve better detection results and effectively improve the single detection area. From Fig. 9, it can be observed that the detection results of the large field of view camera have obvious stretching deformation, and our GelFinger sensor can effectively solve this problem by rotating the camera. On the other hand, using the tactile image acquisition method as shown in Fig. 1, the contact area between the sensor and the object is mainly focused on the middle part of the curved surface, so the area captured using the large field of view camera is not too advantageous compared to the normal field of view camera. The camera inside the proposed sensor rotates while the robot arm makes a small movement, similar to the movement of a human touching an object, so a larger area can be detected. Additionally, compared to the flat silicone detection surface type of sensor, the GelFinger sensor has more advantages in surface detection applications and can effectively expand the scope of practical applications.

## VI. PIPELINE CRACK DETECTION

To demonstrate the practical utility of our proposed GelFinger sensor, we conducted a pipeline crack detection task as an illustrative example. Pipeline cracks can be challenging to detect using conventional visual inspection methods, as they are often located in hard-to-reach areas and may not be visible to cameras. By leveraging the high-resolution tactile imaging capabilities of the GelFinger sensor, we were able to detect the cracks with a high degree of accuracy.

### *A. DeepLabv3+ Algorithm*

The DeepLabv3+ model is widely regarded as the most advanced deep learning model for semantic segmentation, owing to its superior accuracy and computational efficiency. Its encoder-decoder structure, as shown in Fig. 10, has been instrumental in achieving these outcomes. In this paper, we leverage the crack tactile images acquired by GelFinger as the input and use the DeepLabv3+ to identify and segment cracks in the tactile images.

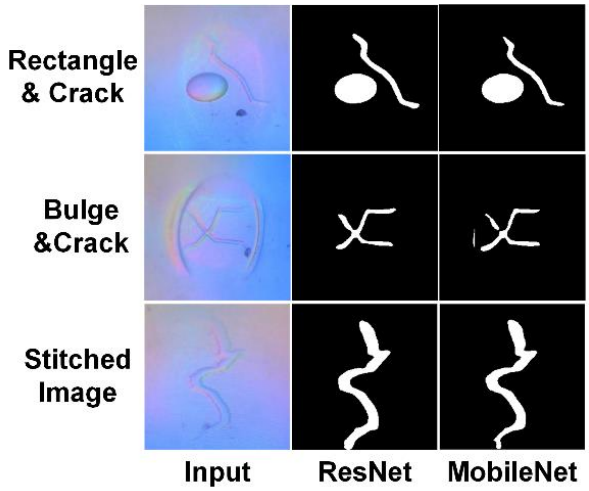


Fig. 11. Visualization of the segmentation results using ResNet and MobileNet as backbone networks, respectively, compared to the input.

### *B. Tactile Dataset and Model Training*

Our experiment involves using the GelFinger sensor to capture tactile images, which serve as the dataset for training and testing our crack detection models. The dataset consists of two parts: a single image captured by the sensor and a complete image stitched from multiple angular images. In total, the dataset contains 1,000 tactile images.

To evaluate the efficacy of our approach, we designed and printed a variety of object models, which can be classified into two types: curved surfaces and flat surfaces. Each model features cracks of varying widths, ranging from 0.5 to 2 mm, with both regular straight lines and irregular intersecting patterns. Additionally, we designed regular shapes such as squares, circles, and triangles to detect other types of curved surfaces or plane defects besides cracks.

We mounted the GelFinger on the UR3 robotic arm, and developed acquisition software to control the arm's movements and capture tactile images along a predetermined trajectory. To train and test our models, we randomly divided the dataset in a ratio of 8:2, with 800 images used for training and 200 images used for testing. We used ResNet and MobileNet as the backbone networks and loaded the pre-trained network structure weights during the training process to accelerate network convergence. The initial learning rate was set to 0.007, and the SGD optimizer was used. The weight decay rate was set to 0.0004, the batch size was set to 8, and the models were trained for 300 epochs.

### *C. Crack Detection Experiments and Results*

The crack detection results are presented in Table III, which shows the Mean Intersection over Union (MIoU) scores for the DeepLabv3+ model using the two different backbone networks. The results indicate that when ResNet is used as the backbone network, the MIoU is higher than that of MobileNet.

We further analyzed the performance of our approach by testing it on three different scenarios: cracks and round defects,

cracks and bulges, and cracks formed by splicing multiple images. Fig. 11 illustrates the tactile images captured by GelFinger and the corresponding crack prediction effect for each scenario. Our results show that the ResNet-based model outperforms MobileNet in terms of accuracy and boundary smoothness. Specifically, ResNet shows smaller errors of the judgment area, and smoother boundaries, especially in the prediction of bulge defects.

## VII. Discussion and Conclusion

In this paper, we propose a novel visual-tactile sensor GelFinger, which is designed to resemble the shape of a human finger with a large contact area. The sensor incorporates a micro motor that drives the camera, enabling multi-angle capture of tactile images. The new design results in higher image quality, and overcomes the problem of severe stretching and deformation of tactile images due to the curved contact surface of the visual-tactile sensor. The APAP image stitching algorithm is employed to combine the multi-angle tactile images, thereby increasing the primary detection area of the sensor. The sensor is used for the pipeline crack detection experiment, in which we trained the DeepLabv3+ model on our tactile image dataset. The experimental results show that our GelFinger sensor achieved accurate and effective detection of pipeline cracks. Additionally, GelFinger can also detect touches on irregular surfaces.

While the GelFinger sensor has shown promising results in detecting pipeline cracks and other surface defects, there is still room for improvement in design and preparation, as well as the flexibility of its use. Considering the problem of uneven brightness at different points inside the sensor, diffusers could be added to the top of the LEDs. Improve the preparation method of the transparent support layer, for example, by pouring clear epoxy resin into finger-shaped molds, to simplify the preparation process and increase the preparation efficiency. There is a lack of an algorithm for detecting contact points or areas. The controls for both the robot arm and the motor are pre-determined, also limiting the range of applications for which the sensor can be used. To address this, we plan to develop a set of control algorithms that will enhance the flexibility of the sensor's use, enabling it to adapt to different types of surfaces and detect defects in various settings. By improving the control system, we believe that the GelFinger sensor can be a valuable tool for a wide range of industrial and research applications.

## References


[1] Y. Wan, Z. Qiu, Y. Hong, Y. Wang, J. Zhang, Q. Liu, Z. Wu, and C. Guo, "A highly sensitive flexible capacitive tactile sensor with sparse and high-aspect-ratio microstructures," *Advanced Electronic Materials.*, vol. 4, no. 4, p. 1700586, March, 2018.

[2] R. S. Dahiya, G. Metta, M. Valle and G. Sandini, "Tactile Sensing—From Humans to Humanoids," *IEEE Transactions on Robotics.*, vol. 26, no. 1, pp. 1-20, Feb. 2010.

[3] R. Romeo, C. Oddo, M. Carrozza, E. Guglielmelli, and L. Zollo, "Slippage Detection with Piezoresistive Tactile Sensors," *Sensors.*, vol. 17, no. 8, pp. 1844, Aug. 2017.

[4] W. Yuan, S. Dong, and E. Adelson, "GelSight: High-Resolution Robot Tactile Sensors for Estimating Geometry and Force," *Sensors.*, vol. 17, no. 12, pp. 2762, Nov. 2017.

[5] N. Lepora, Y. Lin, B. Money-Coomes and J. Lloyd, "DigiTac: A DIGIT-TacTip Hybrid Tactile Sensor for Comparing Low-Cost High-Resolution Robot Touch," *IEEE Robotics and Automation Letters.*, vol. 7, no. 4, pp. 9382-9388, Oct. 2022.

[6] W. Yuan, S. Wang, S. Dong, E. Adelson, "Connecting look and feel: Associating the visual and tactile properties of physical materials." in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2017, pp. 5580-5588.

[7] S. Katyara, F. Ficuciello, T. Teng, F. Chen, B. Siciliano, and D. Caldwell, "Intuitive tasks planning using visuo-tactile perception for human robot cooperation" 2021, *arXiv:2104.00342.*

[8] K. Sato, K. Kamiyama, N. Kawakami and S. Tachi, "Finger-Shaped GelForce: Sensor for Measuring Surface Traction Fields for Robotic Hand," *IEEE Transactions on Haptics.*, vol. 3, no. 1, pp. 37-47, Jan.-March 2010.

[9] S. Dong, W. Yuan and E. H. Adelson, "Improved GelSight tactile sensor for measuring geometry and slip," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, 2017, pp. 137-144.

[10] M. Lambeta, P. Chou, S. Tian, B. Yang, B. Maloon, V. 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, no. 3, pp. 3838-3845, July. 2020.

[11] G. Cao, J. Jiang, C. Lu, D. Gmoes, and S. Luo, "Touchroller: A rolling optical tactile sensor for rapid assessment of large surfaces," *Sensors.*, vol. 23, no. 5, pp. 2661, Feb. 2023.

[12] M. Johnson and E. Adelson, "Retrographic sensing for the measurement of surface texture and shape," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2009, pp. 1070-1077.

[13] M. Johnson, F. Cole, A. Raj, and E. Adelson, "Microgeometry capture using an elastomeric sensor," *ACM Transactions on Graph*., vol. 30, no. 4, pp. 1-8, July. 2011.

[14] R. Li, R. Platt, W. Yuan, A. Pas, N. Roscup, M. Srinivasan, and E. Adelson, "Localization and manipulation of small parts using GelSight tactile sensing," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, 2014, pp. 3988-3993.

[15] D. Gomes, Z. Lin and S. Luo, "GelTip: A Finger-shaped Optical Tactile Sensor for Robotic Manipulation," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, 2020, pp. 9903-9909.

[16] D. Gomes, Z. Lin, and S. Luo, "Blocks world of touch: Exploiting the advantages of all-around finger sensing in robot grasping," *Frontiers in Robotics and AI*., vol. 7, pp. 541661, Nov. 2020.

[17] W. Do and M. Kennedy, "DenseTact: Optical Tactile Sensor for Dense Shape Reconstruction," in *Proc. IEEE Int. Conf. Robot. Automa*., Philadelhia, PA, USA, 2022, pp. 6188-6194.

[18] B. Romero, F. Veiga and E. Adelson, "Soft, Round, High Resolution Tactile Fingertip Sensors for Dexterous Robotic Manipulation," in *Proc. IEEE Int. Conf. Robotics and Automation*., 2020, pp. 4796-4802.

[19] H. Sun, K. Kuchenbecker, and G. Martius, "A soft thumb-sized vision-based sensor with accurate all-round force perception." *Nature Machine Intelligence.,* vol. 4, no. 2, pp. 135–145, Feb. 2022.

[20] Q. Tian, Q. Luo, B. Ge, and Y. Liu, "A methodology framework for retrieval of concrete surface crack′ s image properties based on hybrid model," *Optik*,. vol. 180, pp. 199-214, Feb. 2019.

[21] H. Fu, D. Meng, W. Li, and Y. Wang, "Bridge Crack Semantic Segmentation Based on Improved Deeplabv3+," *Journal of Marine Science and Engineering.*, vol. 9, no. 6, pp. 671, Jun. 2021.

[22] J. Zaragoza, T. Chin, M. Brown, and D. Suter, "As-projective-as-possible image stitching with moving DLT," in *Proc. IEEE Conf. on Comput. Vis. Pattern Recognit.*, 2013, pp. 2339-2346.


IEEE Robotics and Automation Letters (RA-L) paper, presented at ICRA 2024, Yokohama, Japan. Cite as RA-L paper.