Machine-Learning-Assisted Soft Fiber Optic Glove System for Sign Language Recognition

Renjie Zhu, *Graduate Student Member, IEEE*, Dongliang Fan, *Member, IEEE*, Jiang Lin, Huijuan Feng, Hongqiang Wang, *Senior Member, IEEE*, and Jian S Dai, *Fellow, IEEE*

***Abstract*—Sign language recognition devices are effective approaches to breaking the communication barrier between signers and non-signers and exploring human-machine interactions. Wearable gloves have been developed for gesture recognition and virtual reality applications by employing flexible sensors for motion detection and machine learning for data analysis. However, most existing wearable devices present limited sign language translating capacity due to the sensors' design and distribution. Here, we propose a cost-effective dual-hand soft fiber optic glove system consisting of multimode soft liquid-core fiber optic sensors, gyroscopes, wireless printed circuit boards, and batteries for sign language translation. In combination with different deep learning techniques and recognition strategies, the glove system can recognize static gestures and dynamic gestures of American Sign Language, and deduce the meaning of sentences by the sequence of gestures. The soft glove system exhibits a broad sign language range (10 numbers, 26 alphabets, 18 words, and 5 sentences meaning prediction), and high recognition accuracy (98.6% for static gestures, 95% for dynamic gestures). The results also present the recognizing capacity for high-correlated gestures (e.g., "M" and "N"). Finally, we demonstrate its application for controlling the motion of a virtual character through 7 discrete commands in the VR interface.**

***Index Terms*—Soft sensors and actuators, wearable robotics, gesture, posture and facial expressions.**

I. Introduction

THE rapidly developing wearable devices have improved and extended our interaction efficiency and modes with humans and machines through virtual/augmented reality (VR/AR) systems and have demonstrated various promising applications, including motion detection, healthcare monitoring, and haptic interaction, in recent years [1], [2], [3]. Sign language recognition can not only remove the communication barriers between signers and non-signers but extend the interactive experiences of the human-machine interaction (HMI) [4], [5]. Vision-based technologies are popular strategies for hand gesture recognition, relying on different cameras (e.g., RGB cameras, time of flight cameras, and thermal cameras) for data collection. Still, gesture blocking during expression, poor light conditions, and background clutter may cause inaccurate recognition [6], [7]. Compared to vision technology, a smart glove with flexible sensors is a viable solution to hand gesture and sign language recognition due to its real-time responsiveness and portability. Furthermore, applications, including robot-assisted surgery and rehabilitation [8], [9], human-computer interaction, and gesture recognition for communication and entertainment [10], [11], [12], have been demonstrated. Employing flexible sensors on fingers is a general strategy to measure the finger bending state during sign expression, and multiple static signs were recognized by the smart gloves [13]. Signs not only rely on finger motions but the dynamic movement of wrists. Therefore, gyroscopes and internal measurement units (IMUs) were also used to measure the 3D orientation of hands for dynamic sign translating in glove systems [14], [15]. Moreover, a dual-hand glove system can express and recognize more signs and is more applicable for practical applications compared to sign-hand gloves. Still, the existing glove systems suffer limited sign language recognizing capacity due to the distribution (quantity and location) and sensing ability (including response to different external stimuli of fingers and the measurement of position and acceleration during gesture motions) of sensors, which is not applicable to practical scenarios.

Flexible sensors are essential for measuring finger joint bending and finger interactions for sign recognition. However, existing commercial smart gloves rely on multiple inextensible sensors for motion detection, and their large number of precise sensors and calibration during use limit their widespread applications [16]. Soft sensors relying on various principles, including capacitive, resistive, and triboelectric effects, have exhibited great sensitivity to different stimuli [17], [18], [19]. However, capacitive sensors are vulnerable to external electromagnetic interference, and resistive sensors require complex geometries to enhance the sensitivity for bending detection [20], [21], [22]. Triboelectric nanogenerator-based sensors are prevalent in versatile wearable devices due to their dynamic sensing ability and self-powered property. Still, these sensors are limited to detect continuous motion only, due to the instantaneous triboelectric

Manuscript received 10 August 2023; accepted 11 December 2023. Date of publication 26 December 2023; date of current version 8 January 2024. This letter was recommended for publication by Associate Editor S. Song and Editor Y.-L. Park upon evaluation of the reviewers' comments. This work was supported in part by the National Natural Science Foundation of China under Grant 52275021, in part by the National Key R&D Program of China under Grant 2022YFB4701200, in part by the Science, Technology and Innovation Commission of Shenzhen Municipality under Grant ZDSYS20220527171403009, in part by Southern Marine Science and Engineering Guangdong Laboratory (Guangzhou) under Grant K19313901, and in part by the Foundation for Postdoc of China under Grant GZB20230286. *(Renjie Zhu, Dongliang Fan, and Jiang Lin contributed equally to this work.) (Corresponding authors: Hongqiang Wang; Jian S Dai.)*

This work involved human subjects or animals in its research. Approval of all ethical and experimental procedures and protocols was granted by the Southern University of Science and Technology, Human Participants Ethics Committee under Application No. 20210090, and performed in line with the Declaration of Helsinki.

The authors are with the Shenzhen Key Laboratory of Intelligent Robotics and Flexible Manufacturing Systems, Southern University of Science and Technology, Shenzhen 518055, China (e-mail: zhurj2022@mail.sustech.edu.cn; 11930716@mail.sustech.edu.cn; 12332368@mail.sustech.edu.cn; fenghj@sustech.edu.cn; wanghq6@sustech.edu.cn; daijs@sustech.edu.cn).

This letter has supplementary downloadable material available at https://doi.org/10.1109/LRA.2023.3347133, provided by the authors.

Digital Object Identifier 10.1109/LRA.2023.3347133

TABLE I
COMPARISON OF EXISTING GLOVE SYSTEMS (SINGLE GLOVE)

Glove System	Sensor Types (quantity)	Recognizing Capacity	Applications
Ref. [13]	soft resistance sensors (10)	12 static signs	sign language translation
Ref. [14]	flexible sensors (5), accelerometers (3), gyroscopes (2)	22 dynamic signs	sign language translation, HMI
Ref. [15]	fiber Bragg grating sensors (10), IMU (1)	-	VR, gaming
Ref. [24]	TENG stretchable sensors (10)	11 static signs	sign language translation, HMI
Ref. [30]	gyroscopes (18)	10 static signs	robotic teleportation
Our work	flexible fiber optic sensors (5), gyroscope (1)	36 static signs, 18 dynamic signs and 5 sentences	VR, sign language translation, HMI

processes [23], [24]. Besides electric-based sensors, optical sensors have also been employed for mechanical deformation monitoring recently due to their multifunctional sensing capacity and immunity to external electromagnetic interference [25], [26]. However, the rigid component of optical fibers and large external optical source and detection setup is not applicable to wearable devices. Existing soft optical sensors are also confined to bulky structures and high-modulus matrices, which may hinder the comfortable motion of hands [25], [27].

Herein, we employ soft fiber optic sensors containing a slender, soft tube with a liquid core in the glove system. The soft shell and liquid core endow excellent flexibility to the sensors for adhering to various substrates. According to the frustrated total internal reflection principle, these fiber optic sensors show great sensitivity to stretching, bending, and pressing. The multimode sensing capacity enables more interactive modes of the soft glove. A comprehensive glove system includes a hardware component of sensor technology for data acquisition and a software component for data processing. Deep-learning technology initiates new approaches to complicated data analysis using feature extraction from external sensor systems and autonomous learning [23], [24], [28], [29]. Combined with deep learning algorithms, gloves with soft sensors have already exhibited their capacity for multiple gesture recognition with high accuracy and demonstrated applications in HMI and VR interaction [13], [14], [15], [24], [30], as listed in Table I.

In this paper, to improve the sign recognizing capacity, we develop a soft, high-precision, wireless fiber optic glove system (as shown in Fig. 1) consisting of soft fiber optic sensors, printed circuit boards, and deep learning modules. Based on the soft sensors for sign motion acquisition and deep learning for data analysis, this dual-hand glove system successfully recognizes over 50 signs (including 36 static gestures and 18 dynamic gestures) and 5 sentences with high accuracy. Additionally, high-correlated gestures (such as "M" and "N", "6" and "W") can also be recognized with high accuracy. We also demonstrate the application in the VR interface for controlling the motions of a robot.

The main contributions of the work presented herein are as follows: 1) Introducing soft fiber optic sensors to the smart glove system for bending and pressing deformation measurement during sign language expression; 2) Presenting a soft wireless fiber optic glove system for static, dynamic signs and sentence translation with the assistance of machine learning; 3) Demonstrating the applications of the soft fiber optic glove system in VR interface for virtual character controlling.

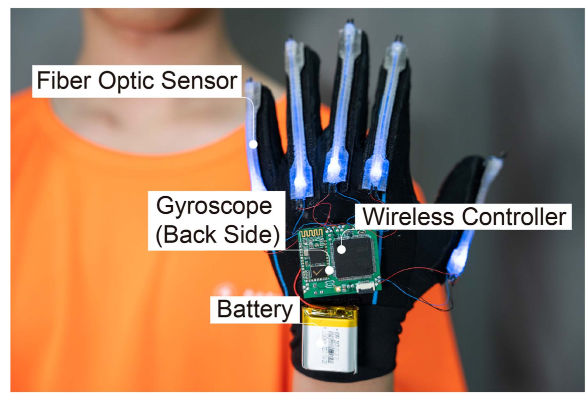

Fig. 1. Photograph showing the implementation of the glove system with fiber optic sensors and a wireless controller board.

II. SOFT FIBER OPTIC SENSOR

A soft sensor with multimode sensing ability is essential to acquiring finger motions in a wireless smart glove system. Considering the appealing characteristics of flexibility, lightweight, high sensitivity, and immunity to electromagnetic interference provided by optical fibers, along with the considerations for comfort and portability of the glove, we develop a soft fiber optic sensor. This section primarily discusses the working principle, design, fabrication, and characterization of the sensor.

A. Working Principle

Most movement and interaction of hand gestures involve finger bending and pressing, as well as changes in hand position. Therefore, a soft sensor with multiple sensing capacities, including stretching, bending, and pressing, is required for hand motion detection. Here, we employ a soft fiber optic sensor containing a soft shell and liquid core for various deformation detection. The total internal reflection occurs in an optical fiber when the refractive index of the inner core is larger than that of cladding with a proper light incident angle. The sensing mechanism of the soft fiber optic sensor obeys frustrated total internal reflection

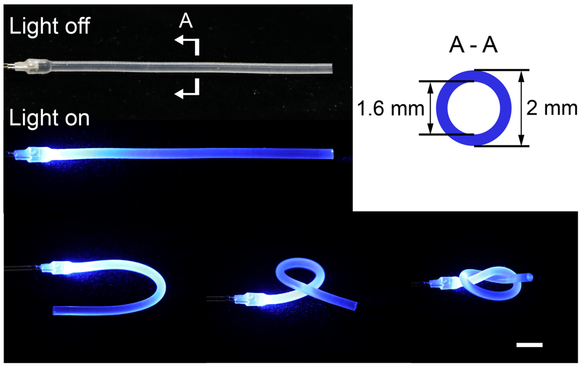

Fig. 2. Soft fiber optic sensor. Scale bar is 5 mm.

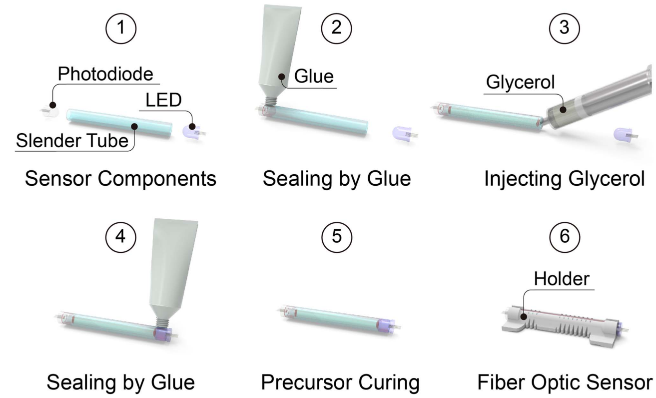

Fig. 3. Fabrication of the fiber optic sensor.

to measure different formations through light intensity. For stretching deformation, the longer pathway of the incident light experiences more attenuation as it propagates through the fiber, which results in a linear decrease in light intensity. The sensor demonstrates its sensing capacity to respond to different external stimuli. In combination with the following deep learning process for the data process, we can distinguish different and combined stimuli. As shown in Fig. 2, when bending or pressing deformation is applied to the soft sensor, the incident angle is smaller than the critical angle at the deformed portion, and a large amount of light escapes from the soft sensor, which results in a sharp decrease in light intensity. By measuring the transmitted light intensity, we can detect different stimuli, including stretching, bending, and pressing. With the photoelectric sensor, we can convert the change in light intensity into a voltage change that is easier to transmit and analyze for subsequent recognition.

B. Design and Fabrication

The fiber optic sensor we proposed is a soft multimode fiber consisting of a liquid core (glycerol, Energy Chemical) with a high refractive index ($n_{core} = 1.47$) and a soft cladding (Ecoflex 50, Smooth-On) with a lower refractive index ($n_{clad} = 1.40$) to guarantee the total internal reflection in the core structure. Both ends of the sensor are capped with a light emitting diode (LED) and a photodetector for light communication, as shown in Fig. 2, and the diameter of the fiber optic sensor was set as 2 mm. The benefits of using liquid core appear in 1) the transparent color, which increases the transporting efficacy of incident light; 2) the soft matrix, which is applicable to conformal substrates; 3) substitutability with other liquid materials for different applications; and 4) biocompatibility for human interaction. In addition, we chose red LEDs (wavelength: 630 nm) for the right-hand glove and blue LEDs (wavelength: 460 nm) for the left-hand glove to demonstrate the design flexibility.

Compared to soft lithography for generating a relatively thick shell with a solid elastic core in a rectangular geometry [25], a highly scalable and effective fabrication technology, named flow casting [31], was employed for the first time to fabricate a soft, slender tube with a thickness of 400 μm as the cladding of the sensor, as shown in Fig. 3. Then, glycerol was injected into the soft tube to form the core-shell structure. A smaller silicone tube (diameter: 2 mm) was set to generate a physical tight connection with LEDs and photodetectors (diameter: 2.5 mm) by inserting LEDs and photodetectors into the tubes. Finally, both ends were sealed by an LED and a photodiode using a small amount of elastomer precursor in case of glycerol leakage. Once the precursor was cured, the fiber optic sensor was produced.

To demonstrate the remarkable compliance and robustness of the fiber optic sensor, a variety of deformations, including bending and knotting, were applied to the soft sensor. Compared to traditional commercial flexible fiber optic sensors, the liquid core with soft cladding structure provides great adaptability and repeatability for various deformations, and the liquid core can be substituted for different applicable scenarios [32]. These advantages make the soft fiber optic sensor a promising candidate, offering improved flexibility, ease of measurement, and robustness for various sensing applications. Further investigation is required to explore the specific sensing characteristics of the fiber optic sensor in the following part.

C. Characterization (Experiments and Simulation)

Herein, we investigated the sensing performance of a soft fiber optic sensor under three deformation modes, namely elongation, bending, and pressing. To quantitatively calculate the optical losses under different forces, we performed solid mechanics and ray optics simulations in COMSOL to simulate the propagation of light in the deformed optic sensor (Fig. 4). Additionally, we defined the output power loss in decibels within the soft sensor (defined as $10log10(I_0/I)$, with I_0 as the output power of the optic sensor without deformation and I as the output power). The experimental results and simulation results confirm each other, and both illustrate the relationship between the light loss rate and the deformation. All experiments were performed in a similar light condition to reduce the effect of ambient light noise.

1) Elongation: The previous works [25] have demonstrated that soft fiber optic sensor exhibits a proportional change in optical power loss with strain since the longer path generates more attenuation. Here, experiments are conducted using a 55 mm long fiber optic sensor with an elongation of 55 mm (i.e., 100% strain). The experimental data is collected using data acquisition equipment (DEWE-43 A, DEWESoft) with a sampling rate of 100 Hz. As the elongation distance increases, the trend of optical power loss is almost linear, yielding a linear, stretch-dependent loss of about 1.46 dB/cm, which is consistent with the simulation model as shown in Fig. 4(a) and (d). Due to the LED light being diffuse, some light leakage occurs, but the majority of the light propagates along the optic fiber successfully. To demonstrate the good repeatability of our soft optic sensor, the sensor signal presents almost the same response under 100 stretch cycles in

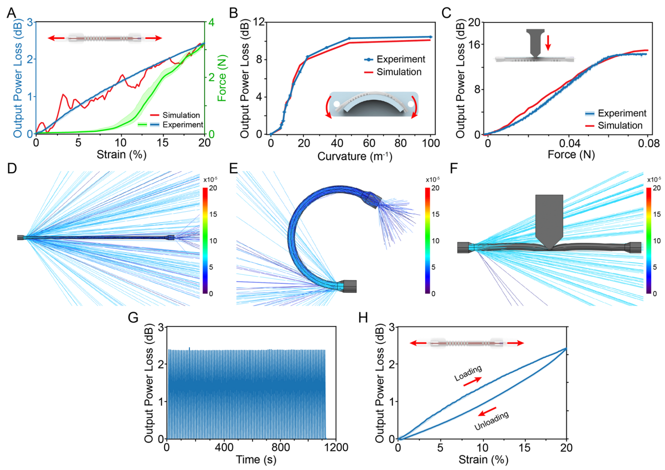

Fig. 4. Characterization of the fiber optic sensor in different deformation modes. Characterization for (a) elongation, (b) bending, and (c) pressing. The propagation of optical rays in the deformed fiber optic sensor (the unit of light power indicated by the color bars for simulation: (d) elongation, (e) bending, and (f) pressing. (g) Stretching cyclic test. (h) Hysteresis of the optic sensor during stretching.

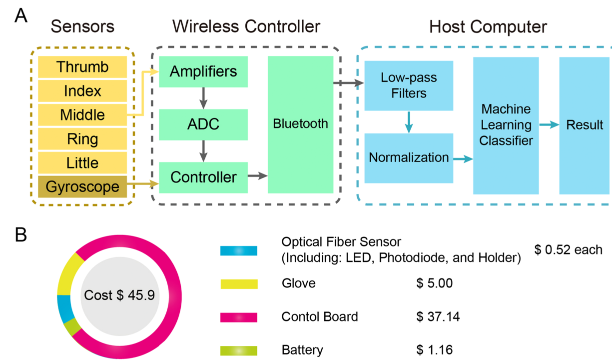

Fig. 5. Circuit diagram showing (a) the signal flow in the glove system from the acquired analog electric signals to the digital signals for gesture recognition. (b) Cost contribution of each component of the soft fiber optic glove system.

Fig. 4(g). We also observed a low hysteresis during the loading and unloading process (Fig. 4(h)).

2) Bending: Bending deformation also causes intensity loss in the sensor. The curvature generated during bending causes a change in the angle of incidence between the light ray and the reflecting plane, resulting in refraction and power loss at the bending position. As shown in Fig. 4(b), we investigated the power output loss of the fiber optic sensor during bending by conducting simulation and experimental tests at a curvature range from 8.33 to 100 m^{-1}. During the bending experiments, the sensor is attached to cylinders with different diameters. When the curvature range is within 8 to 20 m^{-1}, the relationship between output power loss and curvature shows an approximately linear trend. Fig. 4(e) illustrates the process of light propagation along a curved fiber sensor, where the successful transmission of light to the end of the sensor decreases as the curvature increases.

3) Pressing: Due to the low Young's modulus of the elastic material that makes up the fiber, even a small force applied to the fingertip region can cause significant local deformation within the optic fiber, which makes it suitable for pressure sensing. Here, we conducted simulation and experimental pressing tests with a local deformation range set from 0 to 1.8 mm by a pressure head. As shown in Fig. 4(c) and (f), power loss mainly occurs at locations along the optic fiber where the curvature changes dramatically (i.e., at the deformed section under pressure). Furthermore, using a load cell (LSB201, 1 lb, FUTEK), the force applied by the pressure head is also measured.

In summary, the fiber optic sensor is capable of sensing elongation, bending, and pressing and exhibits high repeatability and accuracy, which validate the feasibility of using this sensor in the gloves system.

III. Soft Fiber Optic Glove System

The entire soft fiber optic glove system is a dual-hand smart glove, and each glove consists of three parts: a sensing bock for acquiring hand gesture information, a control block for data collection, and a host computer block for data analysis via machine learning algorithms, as shown in Fig. 5(a). We also exhibit the total price (45.9 dollars) and cost contribution of each component of the glove system (Fig. 5(b)).

A. Sensing Block

The sensing block is employed to gather the topological features exhibited by hand gestures, including the bending motion of fingers and dynamic movement of hands. Therefore, we involved five individual soft fiber optic sensors for bending motion measurement of fingers and one six-axis gyroscope for motion detection of hands in one smart glove. The elastic modulus mismatch between the inextensible fabric glove and the soft sensor is a significant issue in adhering them together conformably, which is critical to measuring the bending gesture accurately. Here, to accommodate the soft fiber optic sensor conformably contact to the curved surface of a finger, we designed and fabricated a dedicated soft holder using 3D printing technology, as shown in Fig. 3. The base contains sawtooth-shaped protrusions that facilitate sensor installation while securely fastening it to the surface of the finger. The soft holder and optical fiber sensor are distributed on the back of each finger and pass through the joint of each finger, which can be well fitted to the finger even when the finger is bent in order to accurately collect topological data of each finger. Before starting the sensing data collection, we calibrate all the sensing modules (such as setting the optical fiber sensor to reach the maximum output power when the finger is stretched).

B. Control Block

As shown in Fig. 1, the wireless controller board is attached to the back of the right hand and powered by a 3.7 V lithium battery (KXW103035, 1200 mAh). The controller block consists of a microcontroller (STM32F407), a six-axis gyroscope (JY901), a Bluetooth module (HC-05), and signal amplifiers (TLC272IDR) connected to optical fiber sensors. The wireless controller uses analog-to-digital converters (ADCs) to receive the voltage signal from the photodetector amplified by the amplifiers, and receives the angular acceleration and angle values from the six-axis gyroscope through the standard universal-asynchronous-receiver/transmitter (UART 2) interface. All collected data is transmitted to the host computer via Bluetooth.

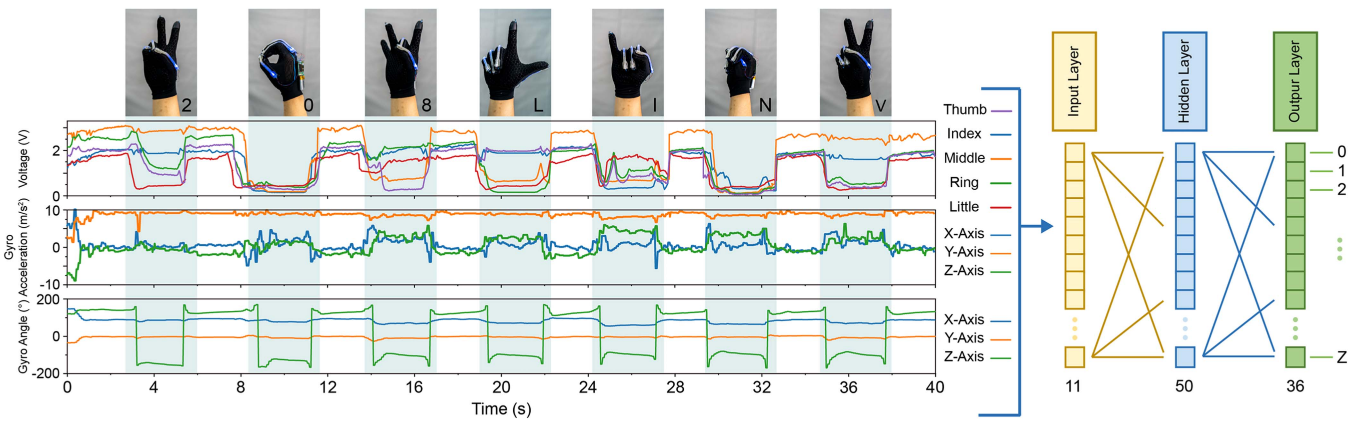

Fig. 6. Demonstration of the glove system, capable of expressing alphabets and numbers with the corresponding generated signal data as recognition patterns and the structure of BP neural network.

C. Host Computer Block

The host computer block is responsible for receiving the data and employing machine learning algorithms for recognition (Fig. 5). It receives wireless data from the glove via Bluetooth, which includes voltage data from the fiber optic sensors and angle values, and angular acceleration from the gyroscope. Prior to processing, the received data undergoes a low-pass filtering process with a cutoff frequency of 80 Hz to eliminate high-frequency noise interference. Then, machine learning algorithms are applied to recognize the data and generate corresponding results. These results can be displayed or used to execute applications on the host computer. The algorithms are described in detail in the following section.

IV. Sign Language Recognition and Applications

A. Static Gesture Recognition

To showcase the real-time static gesture recognition capability, we selected 36 commonly used hand gestures as the fundamental elements of communication from American Sign Language, consisting of 26 alphabets and 10 numbers. Among them, the voltage profiles of 7 hand gestures (“2”, “0”, “8”, “L”, “I”, “N”, and “V”) are shown in Fig. 6. Unlike TENG sensors, which are challenging to recognize discontinuous hand gestures and stationary state of hand [23], fiber optic sensors can recognize hand posture in real-time, even when the hand remains stationary for 4 seconds (Fig. 6). To facilitate the categorization of data under different labels in subsequent steps, we have inserted a reference gesture, the hand naturally extended with the palm perpendicular to the floor, between each static hand gesture, the intentional motion to the right direction caused the rotation angle of Z-axis over 180°, resulting in a jump from 180° to -180° of the gyroscope. In addition, the sampling rate of each ADC is 200 Hz. We used a serial communication protocol to generate 11 data points representing the motion information of a hand simultaneously. Due to inherent restrictions in serial port latency and bandwidth, the transmission rate is 20 Hz. In Fig. 7(a), a matrix is presented that summarizes the correlation coefficients between each gesture and the other 35 gestures. It can be observed that several groups of gestures exhibit strong correlations (indicated by darker colors), which suggests that these gestures (e.g., “2” and “V”, “0” and “O”, “1” and “D”, “M” and “N”, “S” and “T”, “6” and “W”) have high signal similarity

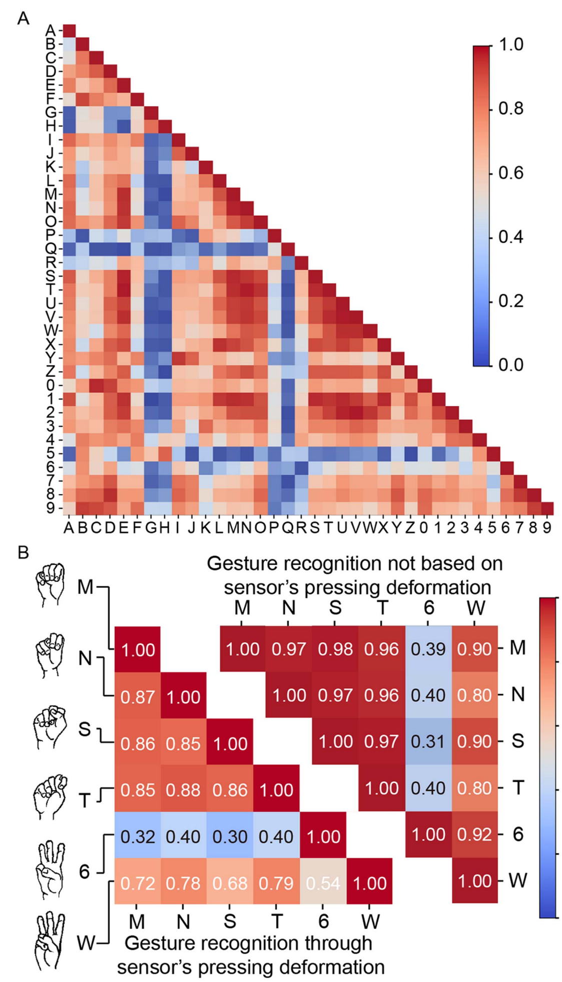

Fig. 7. Correlation coefficient matrix of (a) 36 signs (26 alphabets and 10 numbers), and (b) 6 signs in 3 pairs (“M” and “N”, “S” and “T”, “6” and “W”).

and are, therefore, more likely to be misclassified. Furthermore, for instance, the gestures “M” and “N” exhibit extremely high similarity, differing only in whether the thumb presses down on the index finger. To investigate the impact of pressure on the recognition capabilities of our glove system, we conducted multiple sets of recognition experiments with similar gestures

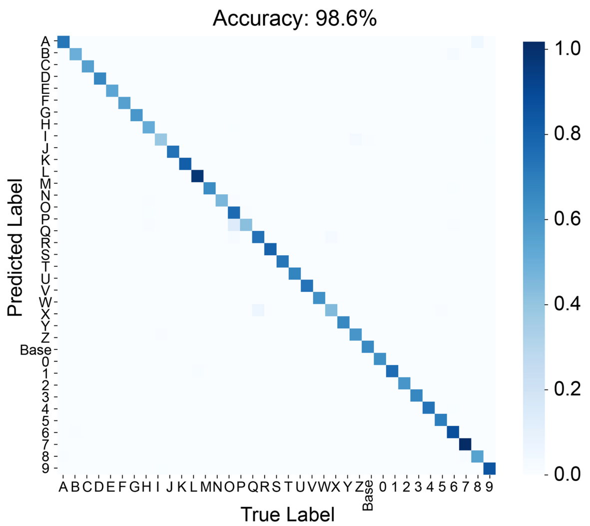

Fig. 8. Confusion map of recognizing 26 alphabets, 10 numbers, and baseline.

using the same glove, and we set expressing gestures without pressure deformation as the control experiments. As shown in Fig. 7(b), this difference may not be detected by a glove system lacking press sensing, resulting in significant confusion between these gestures (with a correlation coefficient close to 0.97). However, the correlation coefficient can be lowered to 0.87 by adding a press-detection sensor. Similarly, the correlation coefficients between "S" and "T" and "6" and "W" decrease from 0.97 and 0.92 to 0.86 and 0.54, respectively, with press detection, leading to improved recognition accuracy and efficacy. Experiments were conducted in accordance with the Declaration of Helsinki. All subjects were informed about the study's general purpose. All trials were approved by the Southern University of Science and Technology, Human Participants Ethics Committee (20210090), and consent was obtained from all participants.

In addition, each static gesture is held for more than 10 seconds (about 10,000 sets of data per gesture) by 3 subjects to ensure sufficient data collection for machine learning model training. Each subject collects data in one session. We calibrate the entire sensing block after we switch the subjects. All data were collected in the order of A-Z, and 0-9. Before feeding the collected data into the neural network, we normalized it to avoid feature weight imbalance and enhance training speed. For recognition and feature extraction, we selected the Back Propagation (BP) neural network with the input layer consisting of 11 neurons to receive 11-channel data (5 optical fiber sensor data and 6 gyro data) and the output layer consisting of 36 neurons corresponding to 36 hand gestures, and a single hidden layer with 50 hidden neurons with the Sigmoid function as the activation function. The collected data are randomly selected into a training set and validation set at a ratio of 9:1. The confusion matrix of the classification is shown in Fig. 8. The horizontal axis represents the predicted value, while the vertical axis represents the actual value. Each gesture has a total number of different data sets (more than 1000 each), which we have trained the network to recognize and present in Fig. 8. 19 of the 36 hand gestures achieved a classification accuracy exceeding 99.5%, and the overall accuracy is 98.6%.

Additionally, we developed a program that integrates the trained BP neural network to enable real-time gesture recognition, as shown in Supplementary Video 1. The 7 hand gestures in Fig. 6 are successfully recognized. Furthermore, some of these gestures are repeated, demonstrating the stability of the static gesture recognition system and its ability to recognize static gestures based on real-time hand gesture data accurately. Note that during the recognition process, the response time in Supplementary Video 1 exceeded 2 seconds, primarily because the recognition program requires manual clicking on the "Recognize" button to start. This button is used to recognize gestures after the hands remain stable to improve recognition accuracy and reduce unnecessary calculations.

B. Dynamic Gesture and Sentence Recognition

For dynamic gesture recognition, we extended a dual-hand glove system and used convolutional neural networks (CNN) for recognition. We first selected 18 gestures and 5 sentences that are daily used in the signer's life for demonstration, as shown in Fig. 9(a). The corresponding signals of these 18 gestures are shown in Fig. 9(b). It includes 10-channel voltage value from fiber optic sensors and 12-channel acceleration and angle data from gyroscopes for both gloves. Unlike the static gesture data discussed earlier, dynamic gestures are continuous, and the relationships between data points are more sophisticated. When we make the same gesture several times to get data, we add an interval gesture, laying hands flat on a table, between each gesture, and then we use this interval gesture to separate the consecutive groups of gestures. For sequence modeling of signals, compared to recurrent neural networks (RNN) and long short-term memory (LSTM), CNN networks can effectively process data with spatial local correlations, such as images, videos, and speech. In other words, CNN is a simple and feasible solution for recognizing time-series signals from fiber optic sensors. We randomly extract the collected data into the training set and verification set in a ratio of 9:1 for training CNN.

The network architecture is shown in Fig. 9(c), and the CNN model contains 3 convolutional layers, with 32 convolutional kernels of size (3, 3) used for each layer, followed by a rectified linear unit (ReLU) as the activation function, which enhances the model's classification capabilities. Max-pooling reduces the number of model parameters by taking the maximum value of the neurons in a local area, which helps to reduce the size of the model and increase the calculation speed. This feature extraction method using stacked multiple convolutional and pooling layers allows the CNN network to gradually understand the complex structure and semantic information of the input data.

The entire CNN network takes 22 (Signal channels) $\times$ 200 (Timesteps, 2 seconds) matrices as input. After extracting enough features, the classification result is calculated through two fully connected layers and a softmax layer. The output of the CNN is the probability that the gesture belongs to 18 different gestures, and the category with the highest probability is our predicted gesture. The prediction results are presented in the form of a confusion matrix in Fig. 10. The horizontal axis represents the predicted value, while the vertical axis represents the actual value. From the Fig. 10, we can obtain that the recognition accuracy is 95%.

To recognize sentences, initially, we split them into individual words and perform word recognition using a simple dynamic window method. Based on the order of the recognized words, the corresponding sentence can be achieved. As shown in Supplementary Video 2, the system accurately recognizes "DIZZY", "ME", "STOMACH", and "HURTS" in sequence and finally

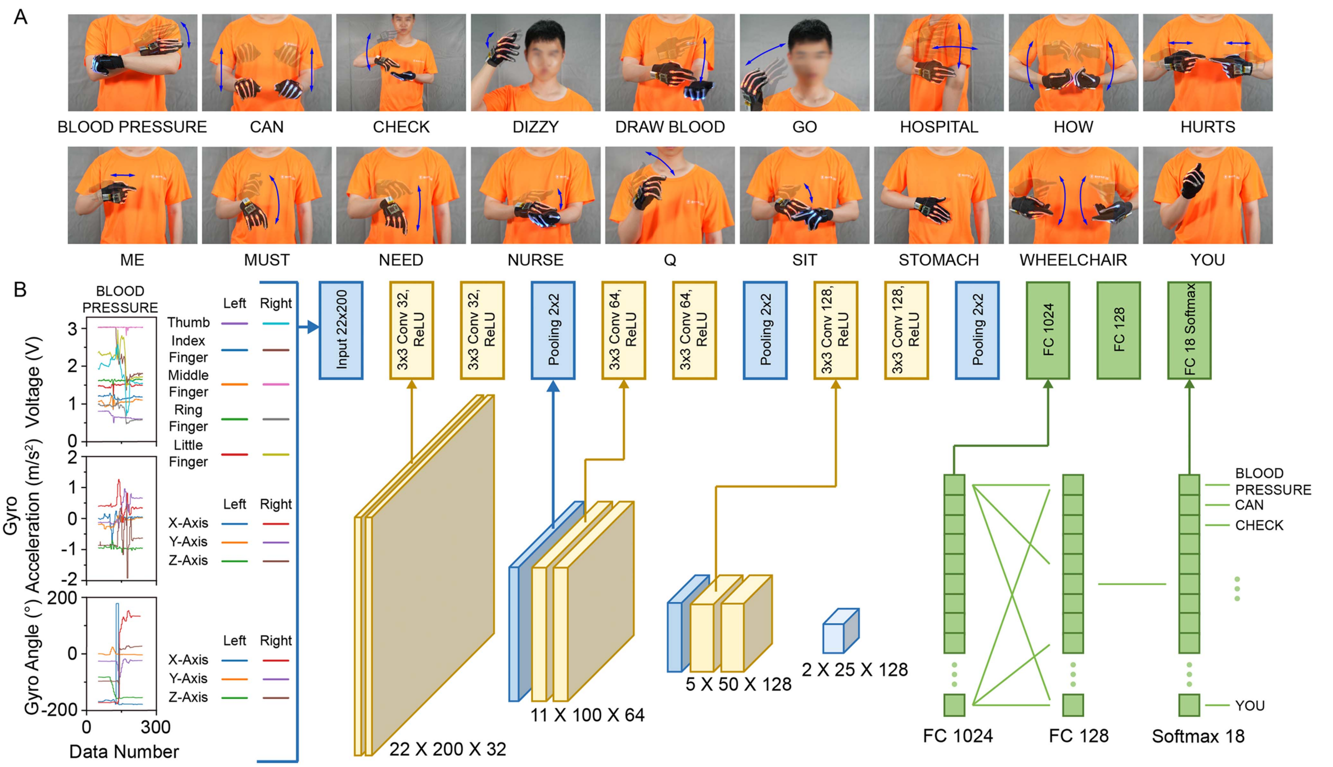

Fig. 9. Demonstration of the glove system, capable of recognizing 18 words. (a) Photographs of 18 words or gestures, in which the translucent and opaque gesture images show the starting and final state of the gesture, respectively. (b) The corresponding generated signal data of the word "BLOOD PRESSURE" and the structure of the CNN network.

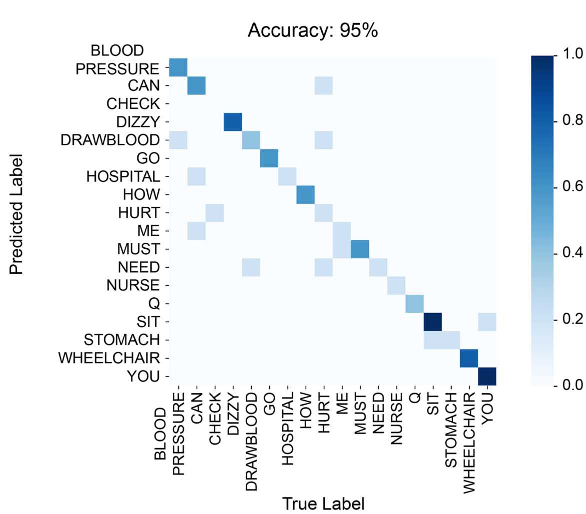

Fig. 10. Confusion map of recognizing 18 words.

translates them into the meaning of "I'm dizzy, and my stomach hurts". Overall, the proposed glove system has great potential in sign language recognition and effectively promotes communication between sign language users and non-sign language users.

C. Virtual Reality Application

Based on the ability of the glove system to recognize both static and dynamic gestures, we explored its potential for virtual reality application. Initially, a set of gestures is designed to control the motion of a virtual character in the Unity interface (version: 2020.3.40f1). Similar to our previous approach, a CNN network is trained to recognize these gestures for control. Therefore, after the subject performs a given hand gesture, the glove system recognizes it and sends the corresponding command to the VR interface to control the motion of the robot. As shown in Supplementary Video 3, we demonstrated the glove system to control a virtual character through various gesture commands, including moving, turning, jumping, rolling, and running. The corresponding signal data during the operation process is shown in Fig. 11. In addition, some of the gestures are repeated to further validate the robustness and accuracy of the recognition system. All executed gesture commands are accurately transmitted, demonstrating the glove system's potential for virtual reality control.

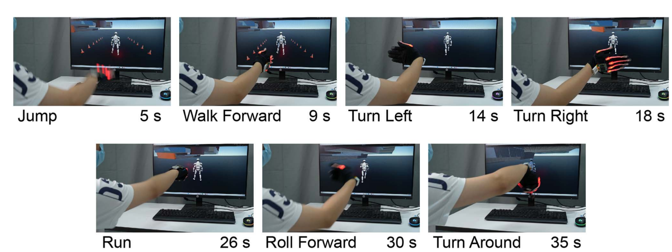

Fig. 11. Demonstration of the glove system for the robot's motion control according to gesture recognition in the VR interface.

V. Conclusion

In this letter, we proposed a dual-hand soft fiber optic glove system consisting of soft fiber optic sensors, six-axial gyroscopes, printed circuit boards, and batteries, which enables robust real-time tracking and recognition of hand movements. The liquid-core soft optic sensor was capable of detecting stretching, bending, and pressing deformations, which is applicable for finger movement and interactions. In combination with gyroscopes for dynamic motion measuring, the sensing system can measure versatile static and dynamic hand gestures. The machine learning technology was involved in data processing, and the smart glove system achieved high accuracy in recognizing 10 digits, 26 alphabets, 18 words, and 5 sentences using this glove system (98.6% accuracy for static gestures and 95% accuracy for dynamic gestures). Furthermore, to explore the potential in VR applications, we demonstrated gesture-based control of the motions of virtual characters in a VR interface.

Future works include improving the sensing capacity of soft sensors, i.e., measuring each joint bending motion of fingers, optimizing the machine learning algorithms for fast recognition, adding haptic feedback devices for improving the interacting experience, configuring continuous recognition windows for continuous recognition, and extending the applications of our smart glove for more practical scenarios between humans and machines. In actual use, we can add a cover to the outside of the light to prevent the interference of light noise. The different colored LED lights in our picture are for better display of gloves and sensors.

References

[1] J. Yin et al., "Wearable soft technologies for haptic sensing and feedback," *Adv. Funct. Mater.*, vol. 31, no. 39, 2021, Art. no. 2007428.

[2] R. Zhu et al., "Soft robots for cluttered environments based on origami anisotropic stiffness structure (OASS) inspired by desert iguana," *Adv. Intell. Syst.*, vol. 5, no. 6, 2023, Art. no. 2200301.

[3] D. Li et al., "Origami-inspired soft twisting actuator," *Soft Robot.*, vol. 10, no. 2, pp. 395–409, 2023.

[4] R. Rastgoo et al., "Sign language recognition: A deep survey," *Expert Syst. Appl.*, vol. 164, 2021, Art. no. 113794.

[5] H. Cooper et al., "Sign language recognition," in *Visual Analysis of Humans: Looking at People*. Berlin, Germany: Springer, 2011, pp. 539–562.

[6] P. K. Pisharady and M. Saerbeck, "Recent methods and databases in vision-based hand gesture recognition: A review," *IEEE Comput. Vis. Image Understanding*, vol. 141, pp. 152–165, Dec. 2015.

[7] M. Oudah et al., "Hand gesture recognition based on computer vision: A review of techniques," *J. Imag.*, vol. 6, no. 8, 2020, Art. no. 73.

[8] R. H. Taylor et al., "Medical robotics and computer-integrated surgery," in *Springer Handbook of Robotics*. Berlin, Germany: Springer, 2016, pp. 1657–1684.

[9] S. Ito et al., "A design of fine motion assist equipment for disabled hand in robotic rehabilitation system," *J. Franklin Inst.*, vol. 348, no. 1, pp. 79–89, 2011.

[10] L. Connelly, Y. Jia, M. L. Toro, M. E. Stoykov, R. V. Kenyon, and D. G. Kamper, "A pneumatic glove and immersive virtual reality environment for hand rehabilitative training after stroke," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 18, no. 5, pp. 551–559, Oct. 2010.

[11] D. Kim, J. Kwon, S. Han, Y. -L. Park, and S. Jo, "Deep full-body motion network for a soft wearable motion sensing suit," *IEEE/ASME Trans. Mechatron.*, vol. 24, no. 1, pp. 56–66, Feb. 2019.

[12] A. Haria et al., "Hand gesture recognition for human computer interaction," *Procedia Comput. Sci.*, vol. 115, pp. 367–374, 2017.

[13] S. Shin et al., "Hand gesture recognition using EGaIn-silicone soft sensors," *Sensors*, vol. 21, no. 9, 2021, Art. no. 3204.

[14] P. Pławiak, T. Sośnicki, M. Niedźwiecki, Z. Tabor, and K. Rzecki, "Hand body language gesture recognition based on signals from specialized glove and machine learning algorithms," *IEEE Trans. Ind. Informat.*, vol. 12, no. 3, pp. 1104–1113, Jun. 2016.

[15] C. K. Jha, K. Gajapure, and A. L. Chakraborty, "Design and evaluation of an FBG sensor-based glove to simultaneously monitor flexure of ten finger joints," *IEEE Sens. J.*, vol. 21, no. 6, pp. 7620–7630, Mar. 2021.

[16] M. Caeiro-Rodríguez et al., "A systematic review of commercial smart gloves: Current status and applications," *Sensors*, vol. 21, no. 8, 2021, Art. no. 2667.

[17] H. Phung, P. T. Hoang, H. Jung, T. D. Nguyen, C. T. Nguyen, and H. R. Choi, "Haptic display responsive to touch driven by soft actuator and soft sensor," *IEEE/ASME Trans. Mechatron.*, vol. 26, no. 5, pp. 2495–2505, Oct. 2021.

[18] H. Liu and J. Dai, "An approach to carton-folding trajectory planning using dual robotic fingers," *Robot. Auton. Syst.*, vol. 42, no. 1, pp. 47–63, 2003.

[19] A. Ghafoor et al., "Stiffness modeling of the soft-finger contact in robotic grasping," *J. Mech. Des.*, vol. 126, no. 4, pp. 646–656, 2004.

[20] D. Fan et al., "Self-shrinking soft demoulding for complex high-aspect-ratio microchannels," *Nature Commun.*, vol. 13, no. 1, 2022, Art. no. 5083.

[21] R. Kang, Y. Guo, L. Chen, D. T. Branson, and J. S. Dai, "Design of a pneumatic muscle based continuum robot with embedded tendons," *IEEE/ASME Trans. Mechatronics*, vol. 22, no. 2, pp. 751–761, Apr. 2017.

[22] D. Fan et al., "Innervation of sensing microchannels for three-dimensional stimuli perception," *Adv. Mater. Technol.*, vol. 8, no. 7, 2023, Art. no. 2300185.

[23] F. Wen et al., "AI enabled sign language recognition and VR space bidirectional communication using triboelectric smart glove," *Nature Commun.*, vol. 12, no. 1, 2021, Art. no. 5378.

[24] Z. Zhou et al., "Sign-to-speech translation using machine-learning-assisted stretchable sensor arrays," *Nature Electron.*, vol. 3, no. 9, pp. 571–578, 2020.

[25] H. Zhao et al., "Optoelectronically innervated soft prosthetic hand via stretchable optical waveguides," *Sci. Robot.*, vol. 1, no. 1, 2016, Art. no. eaai7529.

[26] J. Zhou et al., "Conformable and compact multiaxis tactile sensor for human and robotic grasping via anisotropic waveguides," *Adv. Mater. Technol.*, vol. 7, no. 11, 2022, Art. no. 2200595.

[27] R. Islam et al., "Soft robotic hands and tactile sensors for underwater robotics," *Appl. Mech.*, vol. 2, no. 2, pp. 356–382, 2021.

[28] S. Hussain, R. Saxena, X. Han, J. A. Khan, and H. Shin, "Hand gesture recognition using deep learning," in *Proc. IEEE Int. SoC Des. Conf.*, 2017, pp. 48–49.

[29] W. Zhang, J. Wang, and F. Lan, "Dynamic hand gesture recognition based on short-term sampling neural networks," *IEEE/CAA J. Automatica Sinica*, vol. 8, no. 1, pp. 110–120, Jan. 2021.

[30] D. Lu, Y. Yu, and H. Liu, "Gesture recognition using data glove: An extreme learning machine method," in *Proc. IEEE Int. Conf. Robot. Biomimetics*, 2016, pp. 1349–1354.

[31] D. Fan et al., "Flow casting soft shells with geometrical complexity and multifunctionality," *Adv. Mater. Technol.*, vol. 8, no. 8, 2023, Art. no. 2201640.

[32] O. Schueller et al., "Fabrication of liquid–core waveguides by soft lithography," *Adv. Mater.*, vol. 11, no. 1, pp. 37–41, 1999.