Sensor systems

DiGeTac Unit for Multimodal Communication in Human–Robot Interaction

Gorkem Anil Al[1,2,3] (ID) and Uriel Martinez-Hernandez[1,2,3] (ID)

[1] Department of Electronic and Electrical Engineering, the Centre for Autonomous Robots (CENTAUR), University of Bath, BA2 7AY Bath, U.K.

[2] Multimodal Interaction and Robot Active Perception (inte-R-action) Lab, University of Bath, BA2 7AY Bath, U.K.

[3] Center for Digital Manufacturing and Design (dMaDe), University of Bath, BA2 7AY Bath, U.K.

Abstract—Sensing technologies are required to enhance the safety and interaction between humans and robots in the concept of human–robot collaboration (HRC). In this letter, a multimodal sensing unit, i.e., DiGeTac, composed of two layers with distance, gesture, and tactile elements, is presented. Distance and gesture elements are placed on the top layer of the DiGeTac, while the tactile element is placed on the bottom layer. The sensing unit DiGeTac and its data acquisition board together are designed as a sensing module. The sensor performance for detection of distance, hand gestures and touch perception are analyzed with offline and real-time experiments. The sensing module recognizes hand gestures with 96% accuracy and contact location with 88% accuracy using an artificial neural network and a convolutional neural network, respectively. The sensing capabilities of the proposed module are validated in real-time with a collaborative task using the Universal Robot arm. The sensing capabilities of the proposed DiGeTac, as demonstrated in real-time collaborative tasks with the robot arm, highlight its potential to enhance the interaction in HRC settings by effectively recognizing and responding to human commands.
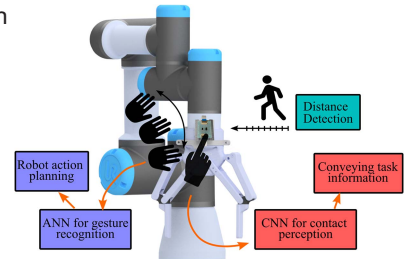
Index Terms—Sensor systems, contactless gesture interaction, human–robot interaction, multimodal sensing interface, touch-based interaction.

I. INTRODUCTION

Traditional manufacturing methods composed of automation systems and robots are not sufficient enough to meet flexible production requirements of mass customization. Thus, human–robot collaboration (HRC) is proposed as a solution for flexible and dynamic production environment where HRC takes the advantage of cognitive abilities of humans, and strength, speed, and accuracy of robots. In the HRC concept, robots should interact with humans and other robots safely while executing collaborative tasks [1], [2]. Therefore, different methods for human safety and interaction have been investigated using sensing devices, e.g., vision, touch, distance, microphones, and wearables [3], [4], [5], [6].

Robots can perceive external changes in their surrounding environment with extrinsic and intrinsic sensing systems. Extrinsic systems are configured around the robot while intrinsic systems are placed on the robot limbs [7]. Examples of extrinsic sensing systems include cameras, radar, and LiDAR sensors, which are highly preferred to develop interaction and safety methods [8], [9]. These sensing systems need to be placed and calibrated in the robot work environment. This process can be time consuming and challenging when the production lines change for manufacturing of different products. In addition, extrinsic sensing systems are also limited by occlusion issues and can detect other objects in static zones. In contrast, intrinsic sensing systems, such as ranging and tactile sensors, offer dynamic motion without restricting the human and the robot motion, making them more

appropriate for close HRC tasks [10], [11]. However, in the literature, a few tactile sensors are capable of providing information in multiple formats to enhance human and robot performance in terms of safety and interaction [12], [13]. Although these multimodal sensors can be operated for close proximity tasks, they cannot be implemented for collaborative tasks where the robot needs full speed operation. This is because these sensors do not include long distance sensing functionality, which limits their application range.

This work presents the DiGeTac unit, which is a multimodal sensor designed with gesture, distance, and a custom-made tactile sensor for industrial HRC tasks. The gesture and distance sensors are placed on the top surface of the DiGeTac unit, while the tactile sensor is designed and placed on the bottom surface. The aim of this multimodal sensing unit is to design a system capable of providing multimodal safety and interaction for collaborative tasks that demand close proximity and long distance perception. Unlike the existing multimodal sensors designed for HRC [12], [13], the DiGeTac offers more interaction and safety features, making it suitable to a wide range of HRC scenarios.

The performance of the multimodal sensor is validated with the following experiments. First, the measurement accuracy of the distance sensor is analysed systematically collecting data with a robot arm. Then, the gesture sensor's hand recognition performance is evaluated employing an artificial neural network (ANN) method. The contact estimation capability of the tactile sensor is tested using a convolutional neural network (CNN) method. Finally, touch-based interaction from the tactile device and contactless interaction, designed with distance and gesture sensors, are validated with HRC scenarios where the human commands the robot to move to specific positions. Overall, the proposed DiGeTac unit can be employed for complex collaborative

Table 1. Technical Specifications of the Sensing Elements Integrated in the DiGeTac Unit

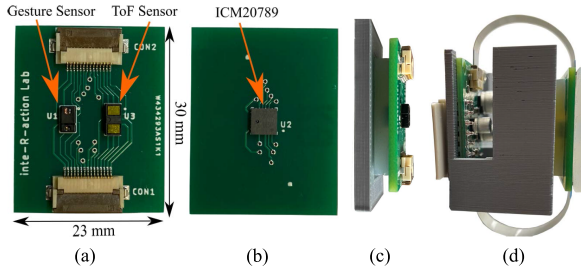| Sensor	VL53L1x	APDS-9960	ICM-20789
Sensing Modality	distance	gesture distance	Three-axis acceleration three-axis gyroscope pressure
Chip size (mm)	4.9 x 2.5 x 1.56	3.94 x 2.36 x 1.35	4 x 4 x 1.4
Range	400 cm	6 gestures 20 cm	$\pm 2/4/8/16g$ $\pm 250/500/1000/2000 dps$ 30-110 kPa
Data Resolution	16 Bit	8 Bit (distance)	16 Bit IMU 24 Bit Pressure Sensor
Frequency	50 Hz	Customisable	10 - 135 Hz

Fig. 1. (a) Distance and gesture sensors placed on the top surface of the sensing unit. (b) ICM-20789 placed on the bottom surface of the unit. (c) ICM-20789 covered with a soft rectangular material and placed on a plastic base to form the sandwich structure of DiGeTac. (d) Sensing module designed placing the DiGeTac and the data acquisition board.

tasks that require different sensing formats, such as in assembly, screwing, and handover.

II. METHODS

A. Sensing Module Design

The DiGeTac unit employs a VL53L1X sensor for distance measurement, an APDS-9960 sensor for hand gesture recognition, and an ICM-20789 chip for tactile sensing. The VL53L1X is a long-range sensor with time-of-flight technology. The sensor can measure a target object up to 4 m. The APDS-9960 is a gesture sensor composed of one LED and four photodiodes that can be used to detect six hand motions, *up*, *down*, *left*, *right*, *far*, and *near*. The tactile sensor is designed with a seven-axis ICM-20789 device consisting of three-axis accelerometer, three-axis gyroscope and a barometric pressure sensor. Table 1 gives the technical specifications of each sensor.

A printed circuit board (PCB) with dimensions of 30 mm × 23 mm (see Fig. 1) is designed, optimizing its layout according to chip routing and tactile sensor design considerations. Distance and gesture sensors are included for contactless interaction while the tactile sensor is designed for contact-based operations. Thus, the distance and gesture sensors are placed on the top surface of the PCB [Fig. 1(a)], and the ICM-20789 is placed on the bottom surface to form the tactile sensor [see Fig. 1(b)]. The data acquisition board of the sensing unit is designed with the required conditioning electronic circuitry [see Fig. 1(d)]. The sensing unit is connected to the data acquisition board with a flexible flat cable on CON1 and CON2 connectors [see Fig. 1(d)].

The chip ICM-20789 has been used by previous works for tactile sensor design [14], [15], [16], and the tactile sensor presented in [15] has been used for the design of a robotic finger and employed in different applications [17], [18]. The tactile sensor in the DiGeTac unit is designed following the fabrication steps described in [15]. A
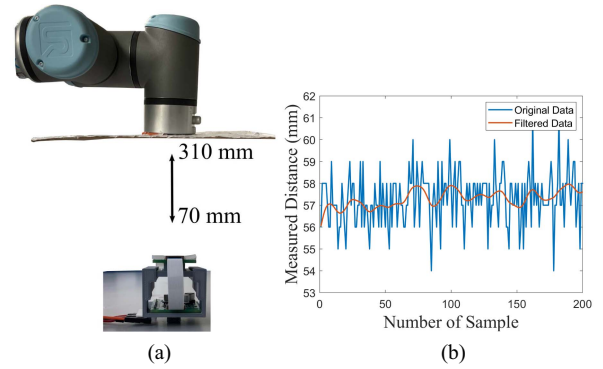

Fig. 2. (a) Data collection from the distance sensor. (b) Distance data measured at 70 mm is filtered with low pass butterworth filter.

soft rectangular case with dimensions of 24 mm × 11.5 mm × 4.5 mm is made from Ecoflex 00-30 silicon material. Then, the soft case is attached to a 3D-printed material to complete the sandwich structure shown in Fig. 1(c). The top surface of the DiGeTac is covered with a soft material layer made of Ecoflex 00-30 [see Fig. 1(d)].

The DiGeTac unit can be attached to the robot surface in various configurations by modifying the connection types to the data acquisition board. A sensing module shown in Fig. 1(d) is designed integrating DiGeTac and the data acquisition board in a plastic component to be attached on the robot.

III. EXPERIMENTS AND RESTULS

A. Evaluation of Sensors Performance

The measurement accuracy of the distance sensor is analyzed using the UR3 robot. The sensing module is placed on the table, and a rectangular cardboard is attached to the flange of the UR3 to cover the view cone area of the distance sensor [see Fig. 2(a)]. First, data from the distance sensor are collected for 10 s at 20 Hz sampling frequency, positioning the UR3 robot 70 mm away from the sensing module. Fig. 2(b) shows that the sensor data contains noise, which makes the sensor output unreliable for robotic applications. The low-pass Butterworth filter is applied to the sensor data to eliminate the noise. The filter is designed as a second order polynomial function to balance the attenuation and delay as follows:

$$y_n = a_0 x_n + a_1 x_{n-1} + a_2 x_{n-2} + b_1 y_{n-1} + b_2 y_{n-2} \quad (1)$$

where y_n is the current filtered output, y_{n-1} and y_{n-2} are the previous filtered outputs, x_n represents the current sensor input, and x_{n-1} and x_{n-2} are the previous sensor input values. In the equation, $a_0, a_1, a_2, b_1,$ and b_2 are the coefficients of the filter calculated using the parameters of 20 Hz sampling rate and 2 Hz cutoff frequency. The second order polynomial function is generated as follows:

$$y_0 = 0.02 x_0 + 0.04 x_1 + 0.02 x_2 + 1.56 y_1 - 0.64 y_2. \quad (2)$$

Fig. 2(b) shows that the filter smooths the sensor signal, and the standard deviation decreases from 1.16 to 0.35 mm. Distance data are collected applying the filter from different target positions to observe the measurement error of the sensor. In the data collection process, the robot is positioned at 70 mm away from the distance sensor and moved upward with 40 mm intervals up to a distance of 310 mm. The standard deviation of each case is less than 0.55 mm, and the mean measurement error from all the cases is 13 mm. In order to compensate the error, 13 mm is added to the distance output.
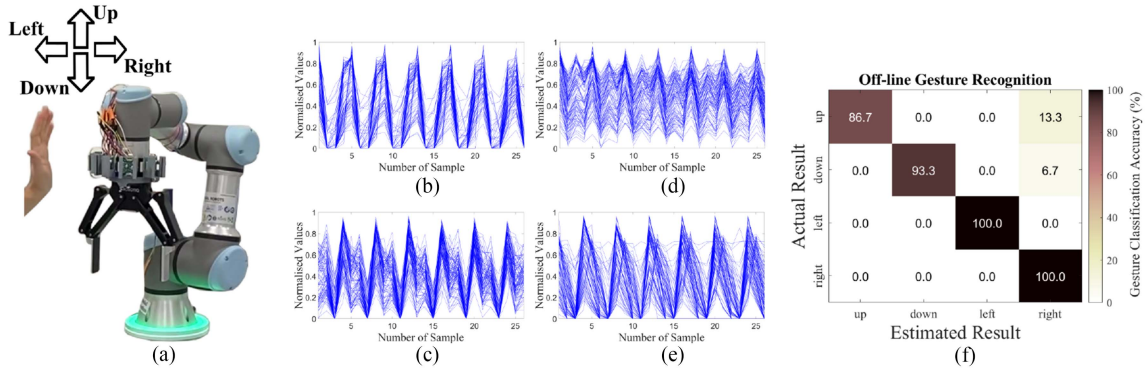
Fig. 3.   (a) Human performs hand gestures. (b) Raw gesture data for up, (c) down, (d) left, and (e) right hand movements. (f) Confusion matrix of the hand gesture recognition in offline mode.
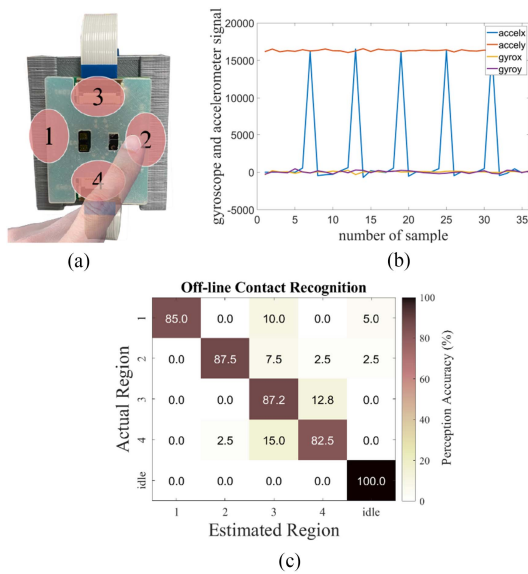

Fig. 4.   (a) Data collection from the four contact locations on the DiGe-Tac. (b) Sample of 4 × 36 matrix used in the CNN method for contact location recognition. (c) Confusion matrix shows the best recognition result achieving 88% accuracy.

The performance of the gesture sensor is evaluated while the sensing module is on the end-effector of the robot arm [see Fig. 3(a)]. The raw data of the gesture sensor, generated by four photodiodes receiving the reflected infrared (IR) energy from the target, are collected at a sampling frequency of 26 Hz while one participant performs *up*, *down*, *left,* and *right* hand gestures in a range of distances between 150 and 200 mm [19]. Each hand gesture is performed 100 times. The number of participants is not critical for this data collection procedure because the different hand anatomy does not affect the collected raw data. Fig. 3(b)–(e) shows the data collected from each hand gesture. These data are used to train an ANN model in MATLAB. The ANN structure, composed of one hidden layer with five neurons, is used for hand gesture recognition [20]. The input dataset is split into 70% training, 15% validation, and 15% testing. The sensing module recognizes hand gestures with 96% accuracy [see Fig. 3(f)].

The tactile sensor in the DiGeTac unit can provide accelerometer, gyroscope, and pressure data. Accelerometer and gyroscope data are used for the recognition of four contact regions (Region 1, Region 2, Region 3, and Region 4) on the sensor [see Fig. 4(a)]. The data

are collected with 12 Hz sampling frequency by pressing on each contact region for 3 s, and this procedure is repeated 200 times for each contact region [19]. Fig. 4(b) shows examples of accelerometer and gyroscope data collected on *x*- and *y*-axes from Region 1. The collected data are used to train a CNN model, which can provide strong recognition capability with kernel-based feature extraction methods. The input data is formed as 4 × 36 matrices composed of 36 samples of the *x*- and *y*-signals from the accelerometer and gyroscope. The input dataset is divided into 80% training, 20% testing. Fivefold validation is implemented and the mean accuracy of 80.4% for contact recognition is achieved. Fig. 4(c) shows the recognition results of the CNN model with the best contact recognition performance with 88% accuracy.

B. Validation of Human–Robot Interaction

Assembly workers often endure nonergonomic positions, leading to workplace injuries, and are susceptible to errors, particularly in scenarios with high product variability [5]. Thus, touch-based and contactless gesture interaction are implemented in a dynamic collaborative task where the human assembles a box putting together small walls. Speed and separation monitoring, which is the collaborative operative mode of the ISO/TS 15066 HRC safety standard [21], is applied with the distance sensor for the safety of the user. Measured distance is implemented dynamically in the speed scaling factor of the robot using real time data exchange method. One sensing module is mounted on the end-effector where the human can use the module comfortably. The distance sensor is calibrated as explained in Section III-A.

The human starts the assembly process by touching the tactile sensor [see Fig. 5(a)]. The robot waits for the human to start the task at the interaction position. When the human presses Region 1 of the DiGeTac, accelerometer and gyroscope data are collected with an Arduino Mega2560 and conveyed to MATLAB. The received data are used in the trained CNN model, and the output is obtained in 3 s, which corresponds to the data collection time. If the CNN model recognizes the contact Region correctly, a message is sent from MAT-LAB to the master device via robot operating system (ROS) to start the collaborative task. After the touch-based interaction procedure, the robot waits for a contactless hand gesture interaction [20]. The human performs one of the four hand gestures [see Fig. 5(b)], and collected gesture data are used in the trained ANN model. The hand gestures *up*, *left*, and *right* are mapped to the position of green, blue, and red side walls, respectively. Collection and processing of gesture data takes 2 s, and after this process, a 2 s wait is added before the robot starts moving. In addition, 2 s are used to allow the human to perform the same action again in case of an incorrect hand gesture
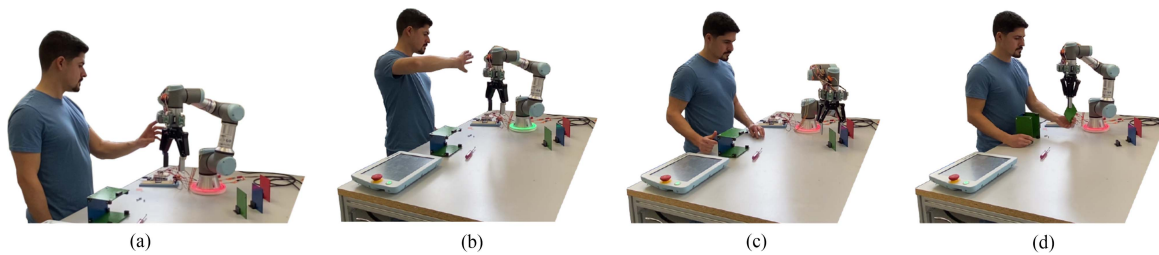
Fig. 5. Steps of the assembly process. (a) Human starts the assembly process by pressing on the tactile sensor. (b) Human performs hand gesture to give a command to the industrial robot. (c) Robot grasps the object asked by the human. (d) Robot handovers the selected object.

recognition. After the incorrect output, if the human holds the hand on the gesture sensor, the ANN method provides an error, and the robot does not perform any action. The recognition output from the model is sent to the master device where the robot actions are planned using MoveIt. According to the recognized hand gesture, the robot moves to the predefined position to grasp the selected object [see Fig. 5(c)]. Then, the robot brings the object to the handover position to deliver the human [see Fig. 5(d)], and the human performs the assembly task.

IV. CONCLUSION

In this letter, the multimodal sensing unit DiGeTac, composed of distance, gesture, and tactile sensors was presented. The unit and its data acquisition board formed the sensing module. Gesture and distance sensors were employed for contactless gesture interaction. An ANN model was trained with the raw data from the gesture sensor, and hand gesture recognition with an accuracy of 96% was achieved. Furthermore, accelerometer and gyroscope data from the tactile sensor were used with a CNN model for touch-based interaction, and the best result for the recognition of contact location with 88% accuracy. These interaction capabilities of the DiGeTac unit were implemented for human–robot interaction in a dynamic assembly task where one user commanded the industrial robot with the unit. While the tactile sensor was used to start the assembly process, contactless gesture interaction was employed to ask the robot to bring the required objects during the assembly process. The experiments showed that different sensing modalities of the DiGeTac can be applied for various use cases in industry. However, during the simultaneous operation of multiple sensors within the DiGeTac unit, a decrease in sensor frequency was observed, which may affect the responsiveness and real-time performance of the system. In our future work, the distance sensor will be used for dynamic speed and separation mode in the assembly task. Also, more participants will be recruited for the HRC task to evaluate the performance of the sensor and collaborative task.

DATA AVAILABILITY

Data created during this research work is openly available from the University of Bath Research Data Archive at https://doi.org/10.15125/BATH-01388

REFERENCES

[1] V. Villani, F. Pini, F. Leali, and C. Secchi, "Survey on human–robot collaboration in industrial settings: Safety, intuitive interfaces and applications," *Mechatronics*, vol. 55, pp. 248–266, 2018.

[2] A. Castro, F. Silva, and V. Santos, "Trends of human-robot collaboration in industry contexts: Handover, learning, and metrics," *Sensors*, vol. 21, no. 12, 2021, Art. no. 4113.

[3] I. Huang and R. Bajcsy, "High resolution soft tactile interface for physical human-robot interaction," in *Proc. IEEE Int. Conf. Robot. Autom.*, 2020, pp. 1705–1711.

[4] P. Gustavsson, A. Syberfeldt, R. Brewster, and L. Wang, "Human-robot collaboration demonstrator combining speech recognition and haptic control," *Procedia CIRP*, vol. 63, pp. 396–401, 2017.

[5] I. El Makrini, K. Merckaert, D. Lefeber, and B. Vanderborght, "Design of a collaborative architecture for human-robot assembly tasks," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, 2017, pp. 1624–1629.

[6] J. Male and U. Martinez-Hernandez, "Deep learning based robot cognitive architecture for collaborative assembly tasks," *Robot. Comput.- Integr. Manuf.*, vol. 83, 2023, Art. no. 102572.

[7] S. Kumar, C. Savur, and F. Sahin, "Survey of human–robot collaboration in industrial settings: Awareness, intelligence, and compliance," *IEEE Trans. Syst., Man, Cybern. Syst.*, vol. 51, no. 1, pp. 280–297, Jan. 2021.

[8] E. Magrini, F. Ferraguti, A. J. Ronga, F. Pini, A. De Luca, and F. Leali, "Human-robot coexistence and interaction in open industrial cells," *Robot. Comput.- Integr. Manuf.*, vol. 61, 2020, Art. no. 101846.

[9] D. Strazdas, J. Hintz, A. Khalifa, A. A. Abdelrahman, T. Hempel, and A. Al-Hamadi, "Robot system assistant (ROSA): Towards intuitive multi-modal and multi-device human-robot interaction," *Sensors*, vol. 22, no. 3, 2022, Art. no. 923.

[10] G. Pang, J. Deng, F. Wang, J. Zhang, Z. Pang, and G. Yang, "Development of flexible robot skin for safe and natural human–robot collaboration," *Micromachines*, vol. 9, no. 11, 2018, Art. no. 576.

[11] S. Kumar, S. Arora, and F. Sahin, "Speed and separation monitoring using on-robot time-of-flight laser-ranging sensor arrays," in *Proc. IEEE 15th Int. Conf. Autom. Sci. Eng.*, 2019, pp. 1684–1691.

[12] P. Mittendorfer and G. Cheng, "Humanoid multimodal tactile-sensing modules," *IEEE Trans. Robot.*, vol. 27, no. 3, pp. 401–410, Jun. 2011.

[13] Y. Zhou, J. Zhao, P. Lu, Z. Wang, and B. He, "Tacsuit: A wearable large-area, bioinspired multi-modal tactile skin for collaborative robots," *IEEE Trans. Ind. Electron.*, vol. 71, no. 2, pp. 1708–1717, Feb. 2024.

[14] G. A. Al and U. Martinez-Hernandez, "A low-cost compact soft tactile sensor with a multimodal chip," in *Proc. 20th Int. Conf. Adv. Robot.*, 2021, pp. 13–18.

[15] G. A. Al and U. Martinez-Hernandez, "Multimodal barometric and inertial measurement unit based tactile sensor for robot control," *IEEE Sensors J.*, vol. 23, no. 3, pp. 1962–1971, Feb. 2023.

[16] U. Martinez-Hernandez and T. Assaf, "Soft tactile sensor with multimodal data processing for texture recognition," *IEEE Sens. Lett.*, vol. 7, no. 8, Aug. 2023, Art. no. 6004704.

[17] G. A. Al, B. E. A. Oussallam, and U. Martinez-Hernandez, "A single-chip multimodal tactile sensor for a robotic gripper," in *Proc. 20th Int. Conf. Adv. Robot.*, 2021, pp. 364–369.

[18] J. Male, G. A. Al, A. Shabani, and U. Martinez-Hernandez, "Multimodal sensor-based human-robot collaboration in assembly tasks," in *Proc. IEEE Int. Conf. Syst., Man, Cybern.*, 2022, pp. 1266–1271.

[19] G. A. Al, "Dataset for "DiGeTac unit for multimodal communication in human-robot interaction," 2024. [online] Available: https://doi.org/10.15125/BATH-01388

[20] G. A. Al, P. Estrela, and U. Martinez-Hernandez, "Towards an intuitive human-robot interaction based on hand gesture recognition and proximity sensors," in *Proc. IEEE Int. Conf. Multisensor Fusion Integration Intell. Syst.*, 2020, pp. 330–335.

[21] I. Standard, "Iso/ts 15066: 2016: Robots and robotic devices–collaborative robots," *Int. Org. Standardization: Geneva, Switzerland*, 2016.