Gesture Based Shopping using Deep Learning

Prof.R.S.Kakade, Mane Namita Balasaheb, Wable Khushi Dilip, Gugale Nisarg Pradeep, Natake Om Bjarang

HOD, Computer Technology , P.D.V.V.P. Institute of Engineering (Polytechnic) Loni, Maharashtra,

Computer Technology , P.D.V.V.P. Institute of Engineering (Polytechnic) Loni, Maharashtra, India

Computer Technology , P.D.V.V.P. Institute of Engineering (Polytechnic) Loni, Maharashtra, India

Computer Technology , P.D.V.V.P. Institute of Engineering (Polytechnic) Loni, Maharashtra, India

Computer Technology , P.D.V.V.P. Institute of Engineering (Polytechnic) Loni, Maharashtra, India

ABSTRACT

Among the most useful body parts of a human being are the hands. Human hands are one of the few appendages that have opposable thumbs that are useful in various tasks such as the making and use of tools. This edge has been essential in the development of the large scale technologies and the growth that we see today. The hands are an intuitive and natural means of communication that is nonverbal in nature. Therefore, the use of hands comes to humans very naturally that can be highly useful in designing tools and other hand held objects. This has led to the designing of various interfaces that have been facilitating the connection between the man and the machine. This leads us to the touch screen devices that are being used for majority of the human machine interactions. But this can be further improved to facilitate the utilization of gestures of the hand to get the interaction. This calls for a fresh approach to e-commerce that incorporates gesture-based engagement. The suggested method uses Convolutional Neural Networks for gesture identification and interaction of the e-commerce software, making it possible to shop using just hand gestures.

Keyword : - *E-commerce, Gesture Recognition, Convolutional Neural Networks.*

1. INTRODUCTION

Implementations that prioritize simplicity and user-friendliness, especially for individuals with impairments, are increasingly sought after, highlighting the pressing need for accessible architecture. While traditional input methods on e-commerce platforms are considered efficient and intuitive for most consumers, touch-based interactions can present challenges for those with disabilities in certain contexts. Hence, the integration of navigational devices responsive to hand gestures becomes crucial, fostering inclusivity and enhancing consumer experiences. This concept of seamless interaction, where users can navigate effortlessly while engaging in other activities, significantly enhances the utility of computing devices.

Hands, being biologically and structurally endowed with remarkable dexterity, serve as the most economical interface, making hand gestures a natural mode of communication. Incorporating hand movements into Human-Computer Interaction (HCI) applications not only enhances natural interaction but also improves user comprehension. Consequently, the adoption of gesture recognition in HCI solutions proves to be promising. The extensive research on gesture-based human-machine interfaces underscores the significance of gestures as fundamental components of intentional action expression.

While the development of data gloves for capturing hand gestures has been explored, their costliness poses a limitation. Thus, recent advancements have focused on affordable imaging devices and image-based approaches for hand motion identification. Hand gestures, with their ability to convey both consistent and dynamic visual actions, offer a versatile means of communication. The increasing research aimed at enhancing communication between humans and robots further emphasizes the relevance of gesture detection technologies.

Recent advancements, particularly in vision-based techniques, have supported the idea of recognizing hand gestures as a valuable interface. Machine vision, due to its accessibility and effectiveness, has been predominantly utilized in research on hand gesture detection. Novel approaches, such as Baohua Qiang's proposed network and Seunghyeok Shin's computational model, demonstrate significant progress in hand tracking and gesture identification, showcasing the practicality and efficacy of these methods.

Furthermore, innovative techniques like the Few Shot training identification method offer swift adaptation and improved performance in hand gesture recognition systems, capitalizing on minimal training data. This approach not only streamlines the training process but also enhances the system's ability to learn from diverse user interactions, contributing to its efficiency and adaptability.

2. LITERATURE SURVEY

Using 3DCNN, Muneer Al-Hammadi explores the potential of hand gesture recognition [1]. The preprocessing stage employed Gaussian selection to normalize the time dimension of hand motion data and the proportions of lengths of identified bodily and facial parts to standardize spatial dimensions. Two distinct methods were used with 3DCNN for deep learning. The first method involved extracting hand gesture characteristics from the full video using a single 3DCNN implementation. The second method entailed training three independent iterations of the 3DCNN architecture to isolate hand motion characteristics at the beginning, middle, and end of the video clip. These regional characteristics were merged before being input into the classifier, and a combination of multilayer Perceptrons, LSTMs, and auto-encoders was used to fuse the features, with a SoftMax active layer utilized for categorization by both methods.

Jiashan Li presents a structure for adaptive action recognition that integrates global gesture action with local finger movement [2]. Experimental findings indicate the successful extraction of important frames from fluid gestures, which replace all fluid gesture panels, thus eliminating redundant information. Additionally, the study introduces an animated gesture classification model, distinguishing between universal movement characteristics of the hands and localized movement characteristics of the digits within the arm, with rotational and translational movements further subdivided. The efficiency of both static and dynamic motion models is demonstrated through experiments.

Jae-Woo Choi introduces a Frequency Modulated Constant Wave transponder with out-of-band detection capabilities [3]. Radar data handling approach suggested by the authors facilitated the creation of three challenging Frequency Modulated Constant Wave gesture recognition datasets. Investigations were conducted to evaluate the performance of various networks, utilizing either images or sequences as input, in identifying the coarse collection. Experimental validation demonstrated the effectiveness of the Transformer encoder-based classification on Frequency Modulated Continuation Wave hand motion data.

Jun Xu proposes an RGB-D inspired identification technique for dynamic and static hand movements [4]. To enhance detection capability for stationary hand gestures, the researchers developed a unique K Curvature Convex Errors Identification methodology, incorporating S curvature and inclination between two fingers and the hand center. Additionally, a modified Bayesian approach morphing technique and a cohesive feature selection method utilizing Euclidean distance among hand joints and shoulder foundation, and maximum likelihood proportions of skeleton characteristics were suggested for recognizing interactive hand gestures.

Jaime E. Lara develops a modular elevated electromyography electrode matrix for capturing hand musculature movement [5]. The architecture, considering factors such as device size, separation, and electrical dispersion parameters, effectively accommodates the unique anatomy of hands and fingers. Cubic support vector machines algorithms were trained using the extensible high-density electromyogram emitter array to verify its effectiveness in categorizing hand movements. The higher electrode surface area enables high-resolution spatio-temporal observations, validated through formal testing of transmission efficiencies and myoelectric potential transmission through muscle tissue.

Yong Wang suggests using a Hybrid Nonlinear Time Morphing radar to identify and distinguish between hands making periodic and changing movements [6]. The operating frequency waveform is extracted from raw hand gesture data, and price bracket mapping and Doppler-time map estimation are performed. The Multimodal Signal Methodology is used to make an educated guess regarding the angle-time representation of various hand movements. Additionally, a hand motion recognition frequency approach and the Fusion Variational Deformation algorithm for classifying gestures are presented.

Abu Saleh Musa Miah develops a Multi-Branch Concentration Driven Graph and Generic Deep Learning framework for hand motion detection based on a sparse dataset [7]. The structure incorporates a fluid graph-based awareness model that integrates geographical, chronological, and learning algorithms data. A generic neural network transforms non-sequential skeletal knowledge into sequence order, and spatial-temporal, temporal-spatial branching is seen as a graph-based deeper neural network division. Positional anchoring is used to construct distinctive identifiers for each location before the first concentration blocks, aiding in progressive data feeding for attentiveness modeling.

Tiantong Wang implements a hand gesture detection functionality using an elastic bracelet equipped with a multi-channel piezoelectric array [8]. The sensitive array's detecting particles and wearability as a wristband accessory distinguish it from alternative methods for gathering gesture signals. Pressure sensors in the bracelet detect movements caused by wrist tissue distortion, which are transformed into impedance variations captured by specialized interpretation circuitry. Triplet connectivity is utilized for intra-day recognition of hand movements due to its exceptional feature learning capabilities.

Hao Zhou proposes and develops a soft bionic arm for prosthesis use that enhances manipulation and gripping capabilities [9]. The 3D-printed homogenous hand features soft location and proximity sensors in its four digits and thumb, enabling precise control of actions and easy switching between them. This research advances grip strength in soft bionic arms, potentially enhancing their use as prosthetics. The hand's mobility aligns with breakthrough pattern classification based user hand connections, facilitating more lifelike control of hand prosthetics.

3. WORKING AND PROCESS

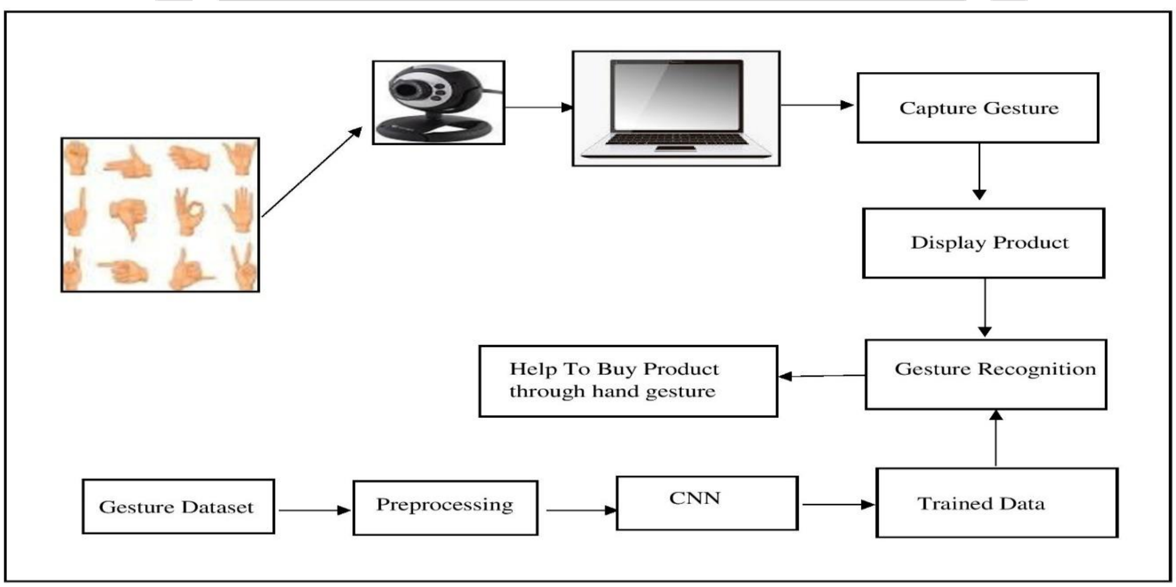

Fig. 1. System Overview Digram

The methodology proposed for gesture-based navigation in a shopping application is realized through deep learning, as illustrated in Figure 1 above, with the steps outlined below.

Step 1: Preprocessing – This initial step involves capturing images of hand gestures using the OpenCV program. Hand motion images are obtained using the VideoCapture feature of the cv2 package. The most common hand gestures for navigation, including left, right, more, next, previous, and purchase, are identified. The YCbCr color space is utilized to detect the skin tone of the arm, followed by removal of the offending region. The image of the hand motion is then resized and transformed into grayscale. The resulting image, scaled down to 96x96 pixels, is stored in a dedicated gesture-only folder. This process is repeated iteratively for each of the 5 gestures until the desired input dataset is obtained.

Step 2: Image Segmentation – Users can view the input dataset obtained up to this point. The dataset, intended for training, is processed by both the training and validation generators. The training generator has a batch size of 64, an image resolution of 96x96 pixels, and a grayscale color palette with classification class settings. Similar features are integrated into the architecture of the validation generator, maintaining a batch size of 64 and grayscale color palette with classification class settings.

Step 3: Convolutional Neural Network (CNN) – This is the core component responsible for identifying and categorizing hand gestures. The input images acquired, preprocessed, and segmented in previous steps are utilized for training the CNN. The input dataset comprises image directories for both training and testing. Each directory is organized into individual subdirectories, representing different hand gestures, along with associated photographs. During training, images are resized to 96x96 pixels. The CNN undergoes training for 500 epochs with a batch size of approximately 64 and a density of 5, reflecting the maximum of five distinct gestures. TensorFlow and Keras libraries in Python facilitate the CNN design. The architecture is depicted in Figure 14.

Following training, a trained model record with the extension .h5 is generated, containing the trained model.

Layer	Activation
CONV 2D 32 X 3 X 3	Relu
CONV 2D 64 X 3 X 3	Relu
MaxPooling2D 2 X 2	
Dropout 0.25	
CONV 2D 128 X 3 X 3	Relu
MaxPooling2D 2 X 2	
CONV 2D 128 X 3 X 3	Relu
MaxPooling2D 2 X 2	
Dropout 0.25	
Flatten	
Dense 1024	Relu
Dropout 0.25	
Dense 5	Softmax
Adam Optimizer	

Fig. 2. CNN Network Architecccture

Step 4: Decision Making – After CNN-based learning, the approach is evaluated for its ability to detect hand gestures related to making a purchase. The OpenCV framework activates video engagement to record hand gestures, which are then cropped and processed. The resized, grayscale images are compared to the components of the .h5 file. Relevant hand gestures are collected and categorized. If the score surpasses a certain criterion, the action is recognized as observed and executed for shopping purposes.

4. APPLICATION

1. Gesture-Based Navigation in Shopping Applications: The primary application is in shopping applications where users can navigate through various products and make purchases using hand gestures. By recognizing gestures such as swiping left or right for browsing products, selecting items, and confirming purchases, users can interact with the application without the need for traditional input devices

2. Accessibility Enhancement:This project can greatly benefit individuals with mobility impairments or disabilities that limit their ability to use traditional input devices. By providing gesture-based navigation, it offers an alternative and more accessible way for these users to interact with shopping applications and make purchases.

3. Hands-Free Interaction: Gesture-based navigation enables hands-free interaction with shopping applications, which can be particularly useful in situations where users have their hands occupied or cannot physically interact with a device. For example, users cooking in the kitchen or performing other tasks can still browse and make purchases using hand gestures without needing to touch a device.

4. Enhanced User Experience: Incorporating gesture-based navigation can enhance the overall user experience of shopping applications by providing a more intuitive and engaging interaction method. Users may find gesture-based navigation more enjoyable and efficient compared to traditional methods, leading to increased engagement and satisfaction with the application.

5. Future Integration with Augmented Reality (AR) or Virtual Reality (VR):This project lays the groundwork for integrating gesture-based navigation with AR or VR technologies in shopping applications. Users could potentially browse and interact with virtual products in a simulated environment using hand gestures, providing a more immersive and realistic shopping experience.

6. Expanded Gesture Recognition Applications: Beyond shopping applications, the methodology developed in this project can be applied to various other domains requiring gesture recognition, such as gaming, healthcare, education, and automotive interfaces. Gesture-based interaction has broad applicability across different industries and use cases, and this project provides a foundation for further exploration and development in these areas.

5. RESULTS

The proposed methodology for achieving effective navigation of an e-commerce application through the use of hand gestures has been effectively portrayed in this research paper. The presented approach allows the use of hand gestures as an input mechanism for controlling and browsing the shopping application. The approach initiates with the gesture dataset being collected for a number of different gestures for the various functions to be performed for the shopping application. These collected gestures are then preprocessed to normalize the images and also resize them to a particular dimension. These preprocessed gestures are then provided to the Convolutional Neural Network for effective training of the model to achieve the trained data. The trained data can now be utilized for navigating the shopping application designed using java. The user can access the system using gestures which are captured. The products are then displayed to the user based on the captured gestures that are then recognized using the trained data. The product can now be purchased through the use of gestures interactions. The approach has been vigorously evaluated for its reliability which has resulted in suitable outcomes.

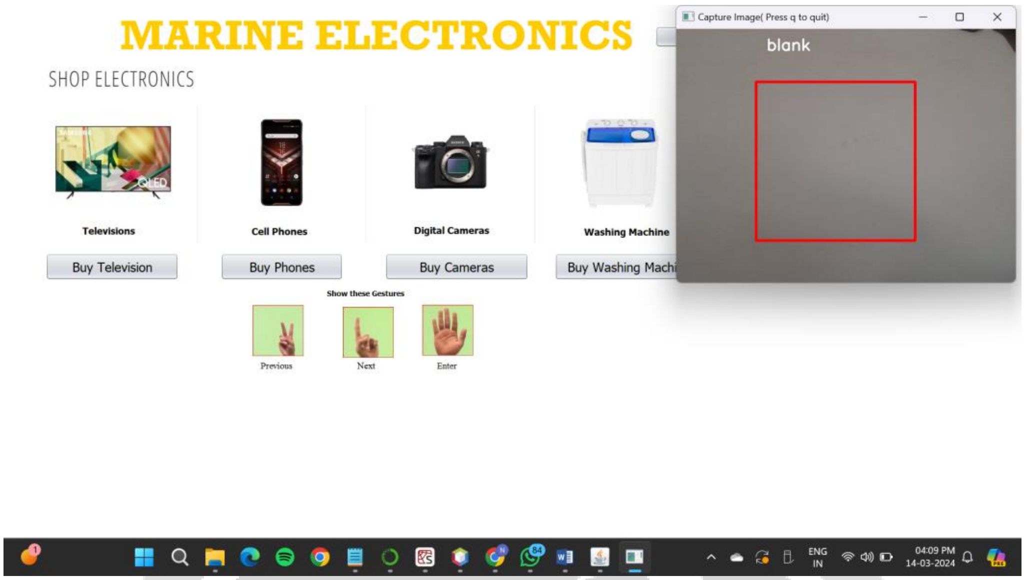

Fig. 3. Empty Gesture

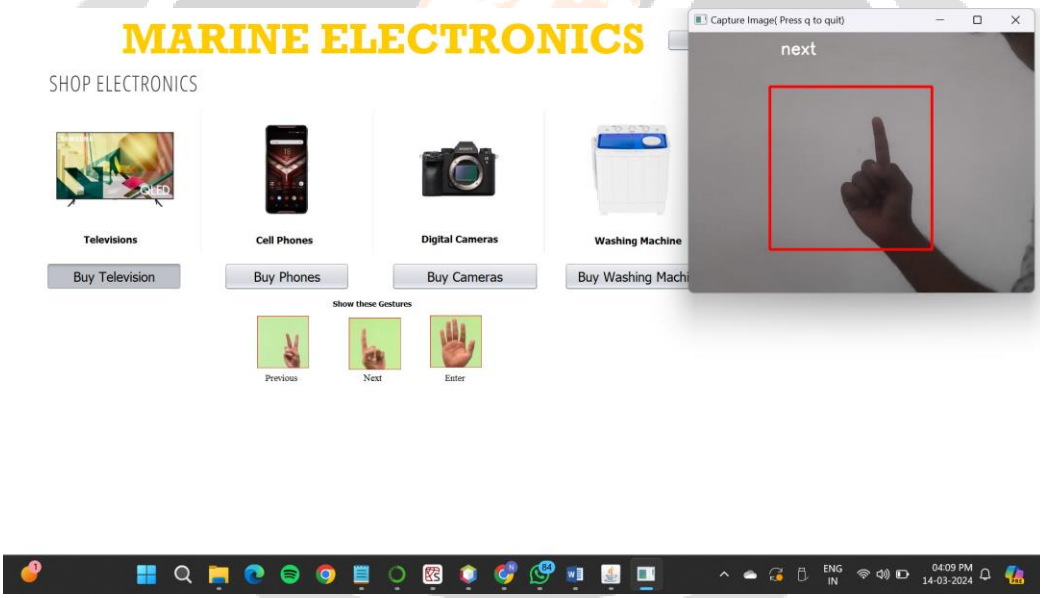

Fig. 4. Next Gesture

6. FUTURE SCOPE

1. Advancements in Gesture UI for Mobile Devices: With the ongoing evolution of smartphone and tablet technology, the future scope involves further enhancements in gesture UI capabilities. This includes the integration of more advanced sensors, such as depth-sensing cameras and infrared sensors, to enable more precise gesture recognition and interaction. Additionally, future mobile devices may incorporate AI-powered algorithms to anticipate user gestures and preferences, leading to a more intuitive and personalized user experience.

2. Expansion of Gesture UI in Gaming:The future of gesture UI in gaming is likely to see widespread adoption and integration into gaming platforms and devices. This may include the development of dedicated gesture-based gaming consoles or peripherals that offer enhanced motion tracking and immersive gaming experiences.

Furthermore, advancements in augmented reality (AR) and virtual reality (VR) technologies may enable gesture-based interactions in virtual gaming environments, further blurring the line between physical and digital interactions.

3. Integration of Gesture UI in Healthcare: The future scope for gesture UI in healthcare is promising, particularly in surgical settings. Further advancements may involve the development of specialized gesture recognition systems tailored to specific medical procedures, enabling surgeons to perform complex tasks with greater precision and efficiency. Additionally, the integration of gesture UI technology into medical devices and equipment could revolutionize patient monitoring and treatment delivery, improving overall healthcare outcomes.

4. Exploration of New Applications and Industries:Beyond the domains of mobile devices, gaming, and healthcare, there are numerous other industries and applications where gesture UI technology could be leveraged. For example, in automotive design, gesture recognition systems could be integrated into vehicles to enable hands-free control of infotainment systems and driver-assistance features. Similarly, in retail environments, gesture UI technology could enhance customer engagement and streamline shopping experiences through interactive displays and kiosks.

5. Research and Development of Advanced Gesture Recognition Algorithms: To fully realize the potential of gesture UI technology, ongoing research and development efforts will be needed to advance gesture recognition algorithms and techniques. This includes the exploration of machine learning and artificial intelligence algorithms to improve the accuracy and reliability of gesture recognition systems, as well as the development of standardized protocols and interfaces to ensure compatibility and interoperability across different devices and platforms.

Overall, the future scope for gesture UI technology is vast and multifaceted, with potential applications spanning various industries and domains. Continued innovation and collaboration among researchers, developers, and industry stakeholders will be essential to unlock the full potential of gesture UI technology and bring about transformative changes in how we interact with digital devices and interfaces.

7. CONCLUSION

In conclusion, the significance of human hands as a natural and intuitive interface for communication and interaction with technology cannot be overstated. From the evolution of opposable thumbs facilitating tool-making to the present-day utilization of hand gestures in human-computer interaction, our hands have played a crucial role in technological advancement. The shift towards gesture-based engagement in e-commerce represents a promising frontier in enhancing user experiences and inclusivity, particularly for individuals with disabilities. Leveraging Convolutional Neural Networks for gesture recognition presents a viable solution to bridge the gap between users and technology, enabling seamless interaction and navigation in computing devices. As research in this field continues to progress, with advancements in imaging devices and machine vision techniques, the potential for enhancing communication between humans and machines through gesture recognition remains vast. Embracing innovative approaches such as Few Shot training identification method further enhances the adaptability and efficiency of hand gesture recognition systems, paving the way for a more intuitive and accessible future in human-computer interaction.

8. REFERENCES

[1] M. Al-Hammadi, G. Muhammad, W. Abdul, M. Alsulaiman, M. A. Bencherif and M. A. Mekhtiche, "Hand Gesture Recognition for Sign Language Using 3DCNN," in IEEE Access, vol. 8, pp. 79491-79509, 2020, doi: 10.1109/ACCESS.2020.2990434.

[2] L. Jiashan and L. Zhonghua, "Dynamic gesture recognition algorithm Combining Global Gesture Motion and Local Finger Motion for interactive teaching," in IEEE Access, doi: 10.1109/ACCESS.2021.3065849.

[3] J. -W. Choi, C. -W. Park and J. -H. Kim, "FMCW Radar-Based Real-Time Hand Gesture Recognition System Capable of Out-of-Distribution Detection," in IEEE Access, vol. 10, pp. 87425-87434, 2022, doi: 10.1109/ACCESS.2022.3200757.

[4] J. Xu, H. Wang, J. Zhang and L. Cai, "Robust Hand Gesture Recognition Based on RGB-D Data for Natural Human–Computer Interaction," in IEEE Access, vol. 10, pp. 54549-54562, 2022, doi: 10.1109/ACCESS.2022.3176717.

[5] J. E. Lara, L. K. Cheng, O. Röhrle and N. Paskaranandavadivel, "Muscle-Specific High-Density Electromyography Arrays for Hand Gesture Classification," in IEEE Transactions on Biomedical Engineering, vol. 69, no. 5, pp. 1758-1766, May 2022, doi: 10.1109/TBME.2021.3131297.

[6] Y. Wang, A. Ren, M. Zhou, W. Wang and X. Yang, "A Novel Detection and Recognition Method for Continuous Hand Gesture Using FMCW Radar," in IEEE Access, vol. 8, pp. 167264-167275, 2020, doi: 10.1109/ACCESS.2020.3023187.

[7] A. S. M. Miah, M. A. M. Hasan and J. Shin, "Dynamic Hand Gesture Recognition Using Multi-Branch Attention Based Graph and General Deep Learning Model," in IEEE Access, vol. 11, pp. 4703-4716, 2023, doi: 10.1109/ACCESS.2023.3235368.

[8] T. Wang, Y. Zhao and Q. Wang, "Hand Gesture Recognition With Flexible Capacitive Wristband Using Triplet Network in Inter-Day Applications," in IEEE Transactions on Neural Systems and Rehabilitation Engineering, vol. 30, pp. 2876-2885, 2022, doi: 10.1109/TNSRE.2022.3212705.

[9] H. Zhou, C. Tawk and G. Alici, "A 3D Printed Soft Robotic Hand With Embedded Soft Sensors for Direct Transition Between Hand Gestures and Improved Grasping Quality and Diversity," in IEEE Transactions on Neural Systems and Rehabilitation Engineering, vol. 30, pp. 550-558, 2022, doi: 10.1109/TNSRE.2022.3156116.