

A Multimodal Human-Machine Interaction System Integrating Sign Language Translation, Emotion Detection and Multilingual Translation

S Prathibah¹, S Prasannikaa², G Rajalakshmi³

^{1,2}*Department of ECE, Sathyabama University*

³*Professor, Department of ECE, Sathyabama University*

Abstract—The technology of human-machine interaction has gained importance in the development of inclusive systems that have helped people who have communication challenges. Individuals who are experiencing problems with speech and hearing tend to experience challenges in everyday communication with the rest of the society, restricting their involvement in social, educational, as well as professional settings. This paper will introduce a hybrid assistive communication system that has been developed to address this disconnect by allowing real time gesture-based interaction with such safety support features. The system translates identified hand gestures into speech and a textual mode thereby enabling the users of the system who use non-verbal mode of communication to communicate with others. Moreover, integrated emergency SOS switch will allow customers to transmit alert messages by means of automated picture snapping and e-mail, which is safer in case of severe circumstances. The system is subjected to Raspberry Pi platform based on the computer vision which provides real-time processing and practical deployment. The given solution shows the possibilities of intelligent assistive technology to enhance the degree of accessibility, independence, and safety of the differently-enabled people.

Index Terms—*Assistive Communication, Human-Machine Interaction, Gesture Recognition, Accessibility Technology, Real-Time Communication, Inclusive Computing, Emergency Alert Systems.*

I. INTRODUCTION

Human machine and interaction have changed a lot over the past few years, and now computers and intelligent systems have the ability to understand human moves and behaviour and make better interpretation on the expression of human beings too. These inventions have

led to space to come up with technologies to improve communication, accessibility and convenience in the day-to-day life. Among the greatest solutions to such development are the assistive technologies that are offered to people with disabilities, especially those that face difficulties in speech and hearing. To such people, their interaction with the immediate environment is often via sign language or other non-verbal forms, which is not necessarily comprehended by the general section of the population. This leads to the creation of communication barriers that often restrict participation in the social sphere and indecency. Intelligent systems that have the ability to decode gestures into comprehensible language can also be important to counter this problem [1].

The deaf and the mute can use sign language as the major means of communication. Though sign languages are rich, expressive and very complex, they do not exist universally to other members outside the deaf community. The absence of this communication easily results in miscommunication and social isolation. The absence of efficient translation systems in educational establishments, workplace and in social places has rendered it quite challenging to allow the non-verbal persons to express their needs or ideas. This has seen scholars working on different computational methods that can help the machine identify the hand gestures and translate them into a spoken or written communication. These advancements prove that computer vision and artificial intelligence have the potential to change assistive communication technologies [2].

Auto recognition Systems Gesture recognition Systems have been extensively researched, as a field of computer vision and pattern recognition. These

systems are normally based on the acquisition of visual data with cameras and the study of the hand features or motion patterns with recognition of the gestures based on a trained model. As machine learning algorithms have improved and more powerful hardware platforms that are affordable are developed, real-time gesture recognition is now more possible. Single-board computers and other embedded computers have low cost and portability, and are now powered with enough processing capability to compute computer vision algorithms. This has allowed the creation of useful assistive technology capable of working beyond the laboratory setting, and assisting in everyday interaction [3].

Although a lot of progress has been made in the research of gesture recognition, most of the existing systems still have a restricted scope of practical use. Other systems do not offer any form of text until gestures have been identified, and, in any case, a third party would have to read and decode the message. Some of them can only work under the controlled laboratory condition and limited vocabulary are used, which makes them less applicable in real life practices. Also [4], some of the research prototypes are based on the high- performance computing resources or cloud-based processing, which may drive up the cost of systems and diminish accessibility. To be really helpful, assistive technologies should be inexpensive, light weight and be able to work in real time on miniaturized hardware platforms.

The other critical process that is usually disregarded in the communication systems is the user safety. Persons with disabilities might find themselves in such situations when they are in dire need of help but they are unable to call out to help by utilizing the conventional means. The implementation of assists like emergency alert systems in the assistive devices will ensure the user remains secure and independent to a great extent. Such systems can add another level of security as well as assurance by allowing users to send automatic alerts or notifications [5] in case of an emergency. Integrative communication support and emergency response services can thus facilitate the establishment of a more holistic support network to the vulnerable users.

The current trends in embedded computing platforms have provided the possibilities of integrating various functionalities in one device. Hardware platforms like Raspberry Pi offer a highly configurable platform on

which computer vision, hardware interface, and network communication modules may be running concurrently. This enables researchers and developers to come up with small systems to incorporate gesture recognition, speech synthesis and emergency alert into a single system. It is specifically appropriate in assistive technology due to its cost and ease of obtaining the hardware.

The system, which is presented in this research paper, is aimed at creating a convenient communication assistant that will translate hand movements into the spoken voice and also be able to issue emergency signals. The gestures of the user are also captured by a camera and then processed by computer vision techniques to detect predefined patterns of gestures. After a gesture is identified, it is placed as a text and processed at the same time into speech using an audio module. This two- output system makes sure that the communication process is possible in both visual and oral forms, which enhance the interaction between the user and other people around him.

Besides gesture translation, the system will have an SOS switch which is used when dealing with emergency cases. Upon activation, the switch also causes the camera to take a shot of the surroundings of the user and sends the image automatically over e-mail to pre-established recipients. This aspect allows the caregivers or relatives to easily interpret the situation to give help when needed. The incorporation of this feature proves that assistive communication technology may also be extended to cover issues of safety.

Inclusive technology is a significant move that the evolution of such systems indicates. Assistive technologies can be used in servicing the needs of individuals with disabilities better by incorporating gesture recognition, speech generation and emergency communication into one device. Moreover, the introduction of these capabilities into the low- cost embedded platform makes the extensive use in education systems, healthcare facilities, and homes more probable.

II. LITERATURE SURVEY

Human-Machine interaction (HMI) has become an important field of research which aims at enhancing human interactions, collaboration, and coordination within human beings and smart computational

systems. As artificial intelligence, machine learning, robotics, and intelligent automation keep improving, the present-day interaction systems are being smarter, more situational, and customer- focused. The early HMI systems were mainly concerned with a graphical user interface and command inputs but in recent times, the concept has been extended to collaborative intelligence where human beings and machines work together to accomplish complicated duties. These systems combine natural language processing system, sensor systems, behavioral analysis and smart decision-making algorithms to facilitate efficient information exchange. This has led to the HMI technologies finding their way into other fields like education, healthcare, industrial automation, transport, and intelligent environment where people have to work hand in hand with their machines to enhance their performance, productivity, and user experiences. Recent studies have investigated how generative artificial intelligence can be integrated in the learning process and result in a better interaction between learners and intelligent systems. Generative AI models can support students in online education platforms, such as real-time responses, personalized feedback as well as interactive mechanisms of learning support, to enhance interactions and learning outcomes [6]. Moreover, emotional intelligence in artificial intelligence has also received much interest and machine learning algorithms have been applied to recognize human affection based on facial expression, speech pattern and behavioral cues to facilitate socially conscious human machine interaction [7]. It is also emphasized in studies that machine learning methods can be used to enhance a human computation interaction, by examining the users input to this interaction and their behavioral pattern to develop dynamic interfaces that react to the user needs and preferences dynamically [8].

The other significance of the interaction system in regard to intelligent interactive system has also grown to be understood as long-term human and robot collaboration. Intelligent continuous learning models enable robots to adapt to the behavior of their users during long durations of time and provide personalized attention and are more accessible to various users [9]. Ergonomics is also carried out to determine the physical efficiency and human factor in the interaction devices of human machine systems and make sure that mechanisms involved in the interaction are

comfortable and reliable in the real world [10]. Moreover, joint decision- making systems that are built on big-data artificial intelligence paradigms have been suggested to facilitate the human decision-making by fuse human and computational intelligence in such a way that systems can process intricate situations and give informed advice [11].

The technologies of human-machine interface are also implemented in other specialized areas including energy system, robotics and neuro-computational interfaces. Interactive control systems have been developed to enhance monitoring and control of distributed energy storage infrastructures so that operators are able to effectively supervise and control the performance of the system by use of smart interfaces [12]. The tools of brain-computer interface offer direct communication channels between neural activities and system of computing in which users can control robots via the electroencephalography (EES) signals [13]. Also, the electromyography signals have been processed using machine learning methods to use them to control the force of the robotic grips, thus enhancing safety and precision of the human-robot collaboration activities [14]. Other machine learning based methods of gesture recognition have also been created to allow robots to interpret dynamic gestures of the human hand allowing human and robot systems to interact in a natural and intuitive way [15].

New developments also underline the usefulness of trust, transparency, and immersive technologies when collaborating with a machine. Autonomous vehicle systems are seeking human-centered artificial intelligence methods to help improve safety and user experience by engaging in adaptive interaction strategies and driver-supervised learning procedures [16]. Other studies have also examined the behavioral cues and explainable AI approaches to approximate the degree of trust when it comes to human- robot partnership so that individuals are guaranteed the safety of engaging with robotic systems in highly challenging workplaces [17]. Besides, machine learning models and benchmark datasets have been proposed to determine movement patterns and assess trust in collaborative robotic systems [18]. Human machine hybrid intelligence systems have been suggested to enhance co-operation between the machine intelligence and human cognition so as to enhance abilities in both directions [19]. Also, to develop immersive environments, digital twin and

virtual reality technologies are becoming more tightly connected with intelligent interaction systems to enable a user to engage with digital models of real-world systems to monitor, train or make decisions about them [20].

III. METHODOLOGY

The approach of the present research is the development of a hybrid human-machine erotic system, which will decode hand signs and offer voice recognition in order to provide an emergency warning system. The system is a combination of computer vision, embedded computing and communication technologies to form a real-time assistive system. There is the structured pipeline of the development process, where image acquisition and preprocessing part are used first, then there is gesture recognition, text conversion, speech generation, and emergency notification. The Raspberry Pi is the central processing unit, and it controls such hardware as the USB camera, the LCD display, the speaker, the LED indicator, and SOS switch. Computer vision and gesture recognition software frameworks are combined so that gestures are correctly detected and identified. The system architecture is such that known gestures will be converted to visual and audio experiences, making users effectively communicate with other people surrounding them.

A. System Architecture and Hardware Setup

The suggested system architecture is created with Raspberry Pi in the centre that is going to be the central controller processing the visual information, running the algorithms of recognition, and providing the output mechanisms control. This hardware developed can be configured with the following devices: USB camera to capture hand gestures, LCD display to show a textual output, speaker to provide audio feedback, LED indicator to give a status alert, and SOS push button to use in case of an emergency event. The power supply will give steady power supply to the embedded components. Video frames are captured by the camera continuously and passed to the Raspberry Pi to be processed. Upon the detection of a gesture, the system translates them into a text message and voice output respectively. The SOS switch will be attached to the Raspberry Pi and its GPIO pins so that the system can detect an emergency input and start the alert processes.

This is a hardware design that is portable, economy-friendly and provides quality real time performance as shown in figure 1.

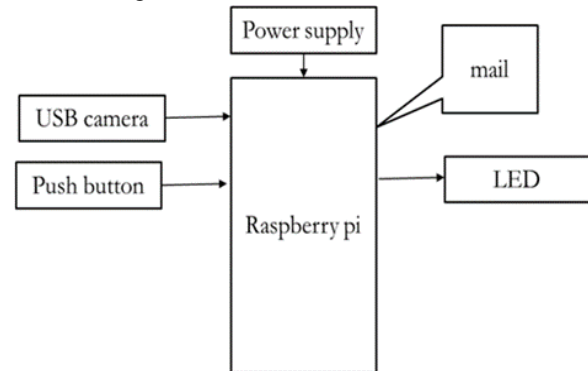


Fig. 1: Block Diagram

B. Image Acquisition and Preprocessing

Gesture recognition pipeline starts with image acquisition in which the USB camera reads an uninterrupted series of frames of the user's hand motions. Computer vision techniques are then used to prepare these frames to be used in detecting and classifying gestures. The first stage is image preprocessing that includes resizing, color conversion and noise suppression to enhance accuracy of recognition. This is done by applying background filtering techniques in order to isolate the hand area with respect to the surrounding environment. An essential set of points in the hand structure is determined with landmark detection techniques that enable the system to recognize the position of fingers and hand pose. This preprocessing step sets it so that the features of the gesture under consideration are well extracted without much interference of the background object or change of light. Improving the image quality and strengthening the aspect of gestures, the stage greatly increases the performance of the next recognition process.

C. Gesture Recognition and Classification

Gesture recognition module tasks are to define the hand gesture which has been captured by the camera and before categorizing it as a specific gesture. The system processes the emerged hand landmarks and pattern of gestures with a trained model of convoluted neural network that is trained to classify gestures. The model checks on spatial features of the hand which encompass finger positioning, angles, and relative relationships among landmarks. Every action is

represented by a pre-programmed command or word in the database of the system. The model during real time operation is recognized by analyzing the framed image appearing and comparing it with the trained gesture patterns to find the most similar one. After an indication is accurately identified, the system recovers the textual message that is linked to the given indication. This classification step is critical in transforming the physical hand movements to decipherable digital instructions that can then be further processed to give out communication.

D. Text Generation and Display Output

Once a gesture has been accurately identified, the system will translate the defined gesture label into a text message that is meaningful. This message is the planned communication of the user and this will be printed on the LCD display device attached to the Raspberry Pi. The text display also gives visual representation of communication which can easily be read by people who communicate with the user. The system makes sure that the characters shown are clear, brief, and coordinated with the understood gesture to make right communication. This is controlled by the Raspberry Pi over the GPIO interface where the text could be dynamically updated along with the changing state of additional gestures. This kind of visual output device proves handy especially in a place where voice communication might not be suitable or where the message has to be confirmed in writing.

E. Speech Generation and Audio Output

The identified gesture can also be translated into an audible speech with the help of a text-to-speech mechanism to make the communication system easier to use. After creation of the textual message, this is absorbed by speech synthesis module, which identifies the text and transforms it into speech. The audio output is sent to a speaker that is connected to the Raspberry Pi. The feature allows the user to interact verbally, in real time, with people around them in order to minimize the use of written form of interaction. The module of speech generation makes sure that the voice output is synchronized with the text recognized and gesture identified. The system is highly effective in terms of accessibility and interaction efficiency given that it is not only visual but is also audible. The synchronization between speech synthesis makes the gesture recognition system a viable communication

aid that should be used on a daily basis.

F. SOS Emergency Alert Mechanism

The system has SOS emergency alert mechanism that is incorporated to increase the level of security of the user in emergency situations. The Raspberry Pi is an emergency trigger that is interlinked with a push button. When the user clicks on the SOS switch, the Raspberry Pi switches on the camera to take an image of the environment around him. This picture, so obtained, is appended to an automated email message, and dispatched to specified emergency contacts. The LED indicator in turn turns on to ensure that the SOS signal has been effectively switched on and sent out. This aspect means that the caregivers or family members will provide the user with timely information on their situation so that they will act on time in case they need any assistance. Having incorporated emergency communication into the assistive device, the system offers communication assistance and individual safety.

IV. RESULT AND DISCUSSION

The deep-sea system that was developed was tested on an experimental level to analyze the real verification of the efficiency of the system with regard to gesture recognition, system response time, reliability of the communication output, and SOS emergency functionality. Lucas was tested using a Raspberry Pi which was wired up with a USB camera, LCD screen, speaker assembly, LED indicator and push button switch. The analysis was based on the relative capability of the system to identify pre-determined gestures and the effectiveness at which they could be transformed into visual and audio outputs in real time. The users were experimenting with various forms of hand gestures in front of the camera and stood within the range of effective operation. The camera took continuous frames of video that were then processed by the system to be able to recognize gestures. Upon the identification of a gesture the message was triggered on the LCD screen and converted to speech using the speaker. The system was stable when gestures were made inside a recognition area and at moderate lighting conditions. The recognition performance was also steady among the various users with a short orientation on gesture placement.

The recognition capability was to be predefined by a set of given gestures that were tested repeatedly to quantitatively analyze the recognition capability. All movements were elementary instructions of communication that were common in everyday speech. The captured hand landmarks were assessed by the system, and those were categorized based on the trained model. Successful recognitions were counted with an aim of establishing the general accuracy. Table 1 shows the results of gesture recognition achieved in case of experimental testing.

Table 1. Gesture Recognition Performance

Gesture Command	Number of Tests	Correct Recognitions	Misclassifications	Accuracy (%)
Hello	30	28	2	93
Help	30	27	3	90
Yes	30	29	1	96
No	30	28	2	93
Thank You	30	27	3	90

Table 1 indicates that the system had a higher recognition of all the gestures that were tested. The recognition accuracy in the gesture Yes was the greatest as the shape was quite distinguishable and had specific finger positioning. Other commands like the help and the thank you were slightly less accurate as they had similarities in the hand orientation and little differences in the manner in which the gestures were done by different users. The errors in the misclassifications usually happened during gestures that were too fast or when some fingers were partly out of the frame of the camera. Although there is this little difference the findings testify that the system can accurately detect usually used gestures in real time communication situations.

The efficiency of the system processing was also observed on top of the accuracy of gesture recognition. Real time communication system demands less time between the detection of gesture and the generation of an output. Thus, the processing latency in response of each of the processing stages was measured to gain insight on the performance of the systems.

Table 2. System Response Time Evaluation

Processing Stage	Description	Average Time (seconds)
Image Capture	Camera captures gesture frame	0.20
Preprocessing	Frame resizing and landmark detection	0.35
Gesture Classification	Model identifies gesture category	0.45
Text Display	Message shown on LCD	0.15
Speech Generation	Text converted to voice output	0.30

Table 2 results demonstrate that gesture classification had the longest processing time due to the fact that the trained model has to go through various landmark features before recognizing the gesture category. The moderate time also demanded in preprocessing through the process of transforming images and getting the features. Nonetheless, display output and generation of speech came right after classification almost instantly. Response time also was not large and handled interactive communication well by letting the user use gestures and get feedback without any significant delay.

Another key aspect that was to be tested in the system testing was the emergency alert feature. The SOS switch was intended to offer direct help during emergency cases. The system used the press of a switch to turn on the camera, take a picture of the environment around, and send the picture in an automated email message to designated people.

Table 3. SOS Alert System Performance

Test Scenario	Button Press Detected	Image Captured	Email Delivered	LED Indicator
Test 1	Yes	Yes	Yes	Activated
Test 2	Yes	Yes	Yes	Activated
Test 3	Yes	Yes	Yes	Activated
Test 4	Yes	Yes	Yes	Activated
Test 5	Yes	Yes	Yes	Activated

Table 3 results show that the SOS mechanism was consistent throughout the testing conditions. The Raspberry Pi was able to receive the signal every time the push button was activated and perform the task of starting the camera and taking an environmental picture, as well as sending the email notification. The LED light also went on immediately and this showed that the alarm process had been triggered. This real time response is essential since it gives the user an assurance that the emergency message has been delivered successfully. The accuracy of this feature is indicative of the fact that it can be useful in helping people who might require quick communication in urgent circumstances.

The entire structure of the made communication system is shown in Figure 2. The diagram emphasizes connecting smart parts of equipment with processing units. The camera records gestures of the hands and transmits image data to the Raspberry Pi. This embedded processor then does preprocessing and gesture classification. Once recognized, the matching message is obtained and sent over two outlets which are visual output in the LCD and voice, via the speaker. The SOS push button in the diagram is also attached to the system via the GPIO pins and thus it activates the emergency alerts.

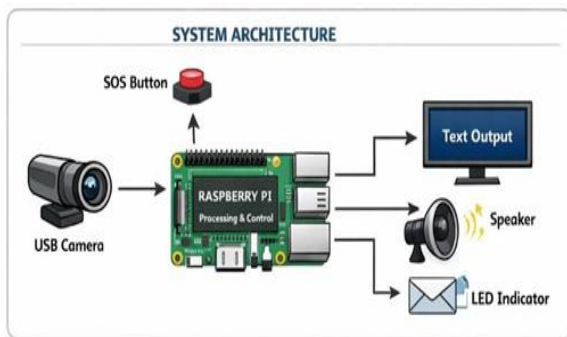


Figure 2. System Architecture of the Proposed Gesture Communication System

This number obviously shows the relationship of the input devices, processing modules, and output components. The architecture makes sure that the gestures are translated into the means of communication and the ability to transmit an emergency alert.

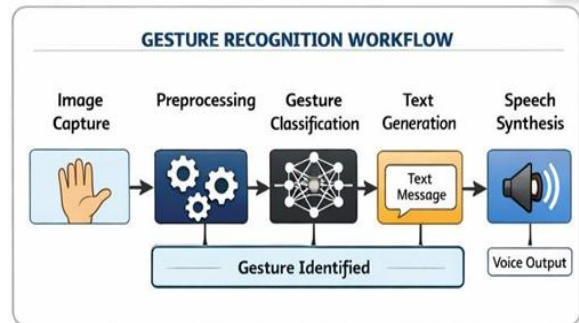


Figure 3. Workflow of Gesture Recognition and Output Generation

Figure 3 illustrates the workflow that makes it possible to turn the non-verbal hand movements into useful communication with the help of the system. The stages are all essential components of the whole system due to its efficiency and accuracy, which makes the results of gesture recognition provided by the system fast and definite.

The way of operation of the SOS emergency alert module incorporated in the system is depicted in Figure 4. With a press of the SOS button, the Raspberry Pi starts a series of actions such as taking of images, creation of notifications, and the activation of LEDs. The image taken is included in an automated mail message and sent over to registered emergency contacts.



Figure 4. SOS Emergency Alert Process

As Figure 4 illustrates, the integration of safety functionality and communication support currently takes place in the system. The automatic alarm system is used so that even the situation when the user is not able to make gestures or speak verbally, the help may be requested. The above feature makes a great contribution to the system in the real-life scenario with regard to its practical usefulness.

Generally, the experimental findings indicate that the

current system was able to integrate gesture recognition, speech generation and emergency notification functions in a small integrated board. The Raspberry Pi offered enough processing power to allow them to do computer vision tasks and yet have a small size and be affordable. Individuals that lack accessibility to communication should also find it easier to communicate when various features are integrated into one device. The system is thus one practical move towards the development of intelligent assistive technology that will aid communication processes and personal safety, which is inclusive.

V. CONCLUSION

This work reported the proposal of a hybrid human-machine interaction system, including a gesture-based communication system and an emergency alert system. The created system allows users to turn hand gestures into a text and speech output, thus making communication devices more effective in the speech or hearing-impaired individuals. The alert feature of SOS is also considered the factor that improves the system since it gives the system an effective safety mechanism where a user will send emergency messages coupled with capturing pictures of the environment. The experimental analysis showed that the system is effective in real-time and gives consistent communication results by way of the visual and audio display. The suggested solution has shown how embedded computing and intelligent interaction technology can be integrated to assistive communication. The future research can be aimed at broadening the vocabulary of gestures, making it more resistant to changing environmental factors and adding more intelligent interaction capabilities.

REFERENCES

- [1] R. Zhang, Y. Qiu, and Y. Li, "An empirical study on human-machine collaborative MOOC learning interaction empowered by generative AI," in *Proc. Int. Symp. Educational Technology (ISET)*, Ho Man Tin, Hong Kong, 2023, pp. 116–120, doi: 10.1109/ISET58841.2023.00031.
- [2] N. Ahmadi and T. Hammond, "Recognizing and responding to human emotions: A survey of artificial emotional intelligence for cooperative social human-machine interactions," in *Proc. IEEE Frontiers in Education Conf. (FIE)*, College Station, TX, USA, 2023, pp. 1–5, doi: 10.1109/FIE58773.2023.10343328.
- [3] F. A. Yaseen, M. S. Ahmed, and M. E. Khudhur, "Enhancing human-computer interaction (HCI) through machine learning application," in *Proc. 8th Int. Symp. Multidisciplinary Studies and Innovative Technologies (ISMSIT)*, Ankara, Türkiye, 2024, pp. 1–4, doi: 10.1109/ISMSIT63511.2024.10757236.
- [4] B. Irfan, N. Churamani, M. Zhao, A. Ayub, and S. Rossi, "Lifelong learning and personalization in long-term human-robot interaction (LEAP-HRI): Overcoming inequalities with adaptation," in *Proc. ACM/IEEE Int. Conf. Human-Robot Interaction (HRI)*, Melbourne, Australia, 2025, pp. 1970–1972, doi: 10.1109/HRI61500.2025.10973812.
- [5] H. Zhang, S. Yang, Y. Wang, and Y. Niu, "An ergonomic study on long-press performance of foot pedals in human-machine interaction applications," in *Proc. Int. Conf. Intelligent Human-Machine Systems and Cybernetics (IHMSC)*, Hangzhou, China, 2025, pp. 88–91, doi: 10.1109/IHMSC66529.2025.00026.
- [6] Z. Liu and Y. Peng, "Human-machine interactive collaborative decision-making based on large model technology: Application scenarios and future developments," in *Proc. Int. Conf. Artificial Intelligence and Digital Technology (ICAIDT)*, Guangzhou, China, 2025, pp. 106–110, doi: 10.1109/ICAIDT66272.2025.00027.
- [7] L. Wei, G. Li, Y. Tao, G. Chen, H. Li, and D. Tao, "Human-computer interaction design of distributed energy storage system," in *Proc. Int. Conf. Artificial Intelligence, Human-Computer Interaction and Robotics (AIHCIR)*, Hong Kong, China, 2024, pp. 187–190, doi: 10.1109/AIHCIR65563.2024.00039.
- [8] H. Bilal, Z. Bi, N. Younis, and H. Abu-Mulaweh, "Human-machine interaction through brain-computer interface," in *Proc. IEEE Int. Conf. Robotics and Biomimetics (ROBIO)*, Chengdu, China, 2025, pp. 2476–2481, doi: 10.1109/ROBIO66223.2025.11376113.
- [9] M. Fahmi, M.-R. S. Noorani, and M. Zakeri, "Machine learning for human-robot interaction using EMG signals to control the grip force," in *Proc. RSI Int. Conf. Robotics and Mechatronics*

- (ICRoM), Tehran, Iran, 2024, pp. 437–442, doi: 10.1109/ICRoM64545.2024.10903539.
- [10] A. Kaur and S. Bansal, “A machine learning-based approach for dynamic hand gesture recognition in human-robot interaction,” in *Proc. Int. Conf. Intelligent Systems and Embedded Design (ISED)*, Rourkela, India, 2024, pp. 1–6, doi: 10.1109/ISED63599.2024.10956785.
- [11] W. Y. Leong, Y. Z. Leong, and W. S. Leong, “Human-machine interaction in the electric vehicle battery industry,” in *Proc. Int. Conf. Applied System Innovation (ICASI)*, Kyoto, Japan, 2024, pp. 69–71, doi: 10.1109/ICASI60819.2024.10547945.
- [12] Y. Wang, S. Debernard, J.-C. Papioul, P. Simon, and J. Floris, “Empowering ADAS with driver-supervised learning of preferences: Parameterization and human-machine interaction,” in *Proc. IEEE Ind. Electronics Society Conf. (IECON)*, Singapore, 2023, pp. 1–6, doi: 10.1109/IECON51785.2023.10312600.
- [13] H. Xu, “Design of human operation management system based on human-computer interaction,” in *Proc. IEEE Int. Conf. Electrical Engineering, Big Data and Algorithms (EEBDA)*, Changchun, China, 2024, pp. 1052–1056, doi: 10.1109/EEBDA60612.2024.10486025.
- [14] C.-G. Xiang and Z. Yu, “Human-machine hybrid augmented intelligence: Human-machine relationship, collaboration and mutual enhancement,” in *Proc. China Automation Congress (CAC)*, Chongqing, China, 2023, pp. 7471–7478, doi: 10.1109/CAC59555.2023.10451218.
- [15] A. L. Florentino and P. T. Aquino-Junior, “Adaptive human-machine interaction for autonomous passenger vehicles: Towards personalization and enhanced user experience,” in *Proc. Brazilian Conf. Robotics (CROS)*, Belo Horizonte, Brazil, 2025, pp. 1–6, doi: 10.1109/CROS66186.2025.11066094.
- [16] A. Kumar, K. Rana, R. Gupta, D. Gunta, A. Mahida, and Sonia, “Human-centered AI for autonomous vehicles: A review of interaction strategies and technologies,” in *Proc. IEEE Delhi Section Flagship Conf. (DELCON)*, New Delhi, India, 2024, pp. 1–6, doi: 10.1109/DELCON64804.2024.10866787.
- [17] L. Huang, Y. Chen, and Y. Zhang, “Research on human-computer interaction-based tool design for agricultural mulch film cleanup assistance,” in *Proc. Int. Conf. Artificial Intelligence, Human-Computer Interaction and Robotics (AIHCIR)*, Tianjin, China, 2023, pp. 526–531, doi: 10.1109/AIHCIR61661.2023.00093.
- [18] G. Campagna, M. Lagomarsino, M. Lorenzini, D. Chrysostomou, M. Rehm, and A. Ajoudani, “Estimating trust in human-robot collaboration through behavioral indicators and explainability,” *IEEE Robotics and Automation Letters*, vol. 10, no. 10, pp. 10218–10225, Oct. 2025, doi: 10.1109/LRA.2025.3600170.
- [19] M. Rehm, K. Hald, and I. Pontikis, “Benchmark movement data set for trust assessment in human-robot collaboration,” in *Proc. ACM/IEEE Int. Conf. Human-Robot Interaction (HRI)*, Boulder, CO, USA, 2024, pp. 934–938.
- [20] C. Zhao, X. Xu, D. Zhang, and Y. Wei, “Intelligent human-machine interaction based on digital twin and virtual reality,” in *Proc. Int. Conf. Robotics, Intelligent Control and Artificial Intelligence (RICAI)*, Hangzhou, China, 2023, pp. 374–378, doi: 10.1109/RICAI60863.2023.10489645.