User-Defined Hand Gesture Interface to Improve User Experience of Learning American Sign Language

Jindi Wang[1 [0000-0002-0901-8587]], Ioannis Ivrissimtzis[1 [0000-0002-3380-1889]], Zhaoxing Li[1 [0000-0003-3560-3461]], Yunzhan Zhou[1 [0000-0003-1676-0015]], and Lei Shi[2 [0000-0001-7119-3207]]

[1] Department of Computer Science, Durham University, Durham, UK
[2] Open Lab, School of Computing, Newcastle University, Newcastle upon Tyne, UK
{jindi.wang, ioannis.ivrissimtzis, zhaoxing.li2, yunzhan.zhou}@durham.ac.uk
lei.shi@newcastle.ac.uk

Abstract. Sign language can make possible effective communication between hearing and deaf-mute people. Despite years of extensive pedagogical research, learning sign language remains a formidable task, with the majority of the current systems relying extensively on online learning resources, presuming that users would regularly access them; yet, this approach can feel monotonous and repetitious. Recently, gamification has been proposed as a solution to the problem, however, the research focus is on game design, rather than user experience design. In this work, we present a system for user-defined interaction for learning static American Sign Language (ASL), supporting gesture recognition for user experience design, and enabling users to actively learn through involvement with user-defined gestures, rather than just passively absorbing knowledge. Early findings from a questionnaire-based survey show that users are more motivated to learn static ASL through user-defined interactions.

Keywords: Human Computer Interaction · Sign Language · User Study

1 Introduction

According to the World Health Organization, around 2.5 billion people will have some degree of hearing loss by 2050[3], and at least 700 million of them will require some kind of hearing rehabilitation. The use of sign language, as well as several other alternative sensory approaches, such as voice reading, writing with the hands, or vibration sensing, are all part of the rehabilitation training courses for people with hearing loss. Although sign language is the most popular means of communication for the deaf, most persons who do not have hearing loss have never taken sign language classes, making communication between these two groups difficult. Thus, in an effort to remove communication barriers between

[3] https://www.who.int/zh/news-room/fact-sheets/detail/deafness-and-hearing-loss

various groups, learning sign language has emerged as a major research area in education.

The majority of the most recent approaches to the teaching of sign languages [9, 10, 16] employ predefined gestures, while approaches based on user-defined interaction technology are more sparse in the literature. For example, Adamo *et al.* [1] proposed the development of a new immersive 3D learning environment to increase the mathematical skills of deaf children by teaching mathematical concepts and American Sign Language (ASL) math terminology through user interaction with fantasy 3D virtual signers and traditional interaction environments. Schioppo *et al.* [20] proposed a sign language recognition method using features extracted from data acquired by a Leap Motion controller from an egocentric view. The method was tested on the 26 letters of the ASL alphabet. In a related development, Phan *et al.* [17] used motion tracking to trial a number of different methods for providing user feedback in a sign language learning system.

Regarding research on the processes by which users can define themselves a vocabulary of hand gestures, Piumsomboon *et al.* [18] conducted research on hand gesture guessability in an Augmented Reality (AR) environment. They invited users to make gestures corresponding to certain tasks, and created user-defined gesture sets to guide the designers in implementing user-centred hand gestures for AR. To the best of our knowledge, there are no studies on learning ASL through user-defined interaction techniques. Hence, the purpose of this paper is to investigate if user-defined interaction techniques can enhance users' learning of ASL. We believe that this is an important research topic since most such systems use hand gestures created by system designers, which do not always reflect user intention.

To accomplish our research goals, we developed a simple system with a user-defined gesture interface for learning static ASL. In the system design, we have identified and taken into account shortcomings of prior systems, including the small data sets used to train the gesture recognizer, the absence of a realistic environment, and most importantly, the user's difficulty in engaging with the system for an extended period of time. With inspiration from Bragg's ASL Sea Battle [5], a sign language game created to help gather user data, we created and integrated a Whack-a-Mole style game with a user-defined hand gesture interface into the system, aiming at boosting user motivation. Finally, we conducted a user study based on a survey designed according to Schrepp's [21] recommendations and concentrated on user experience analysis.

Summarising, the main research question motivating our work, ***"Can user-defined interaction techniques enhance user motivation to learn static ASL?"***, was looked into within the context of a gamified environment for learning static ASL. Our main contributions are as follows:

1. We implemented a user-defined hand gesture interface for ASL learning with a Whack-a-Mole type of game.
2. We conducted a user study to examine if user-defined interaction affected users' experience. The initial results indicate a positive user attitude towards

gamified learning environments and a strong interest of the users in user-defined interactions.

The rest of this paper is organised as follows. The prior work on technology-assisted sign language learning is reviewed in Section 2. The basic design and the features of the proposed sign language learning environment are presented in Section 3. The design of the user study is described in Section 4, while the results are presented and analysed in Section 5. We discuss the main findings in Section 6 and briefly conclude in Section 7.

2 Related Work

The back-end of the proposed system supporting ASL learning with user-defined interaction mainly consists of hand gesture detection and a recognition model. Hence, we review prior research on hand gesture detection and recognition in ASL and user interfaces for creating user-defined hand gestures.

2.1 Sign Language Detection and Recognition

Real-time detection of dynamic hand gestures from video streams is a challenging task since: (i) there is no indication when a hand gesture starts and ends in the video; (ii) a performed hand gesture should only be recognized once; and (iii) the entire system should be designed considering memory and computational power constraints. Bheda *et al.* [3] proposed a method based on deep convolutional neural networks (CNNs) to recognize images of the letters and digits in ASL. Kim *et al.* [13] proposed a novel sign language recognition method, which employs an object detection network for a region of interest segmentation to preprocess the input data. Battistoni *et al.* [2] described a method for ASL alphabet recognition based on CNNs, which allows for monitoring the users' learning progress. Jiang *et al.* [12] proposed a novel fingerspelling identification method for Chinese Sign Language via AlexNet-based transfer learning and evaluated four different methods of transfer learning. Camgoz *et al.* [6] introduced a novel transformer-based architecture that jointly learns Continuous Sign Language Recognition and Translation while being trainable in an end-to-end manner. Zhang *et al.* [25] proposed MediaPipe Hands, a real-time on-device hand tracking pipeline to compute hand landmark positions from a single RGB camera frame for AR/VR applications. Goswami *et al.* [11] created a new dataset for ASL recognition and used it to train a CNN-based model for hand gesture recognition and classification. Finally, Pallavi *et al.* [15] presented a deep learning model based on the YOLOv3 architecture, reporting high recognition rates on the ASL alphabet.

Having reviewed the existing work on sign language recognition, we concluded that Mediapipe is the most suitable tool for the purposes of this paper, and thus, we used it for sign language recognition, benefiting from its highly accurate, real-time detection of hand landmark points. Moreover, as an open-source hand gesture detection framework from Google, it is well-documented and supported.

2.2 User Interfaces for User-defined Hand Gesture

A lot of work has already been done on user-defined hand gesture user interfaces, but most of them support limited functionalities, such as letting the user select one out of two established hand gestures as the one they want to use. For example, Wu *et al.* [24] proposed an interface for users to customize hand gestures and apply them to VR shopping applications in 2019, while they [23] proposed a user-defined hand gesture interface that could be used on in-vehicle information systems in 2020. Besides, conventional means of accessing visual communication markers (VCM) rely on input entry methods that are not directly and intimately tied to expressive nonverbal cues. Koh *et al.* [14] addressed this issue, by facilitating the use of an alternative form of VCM entry: hand gestures. Moreover, to fill this gap Dai *et al.* [7] presented a training system, called CAPG-MYO, for user-defined hand gesture interaction. Takayama *et al.* [22] perform two user studies to derive a user-defined gesture set allowing 13 types of table manipulation.

To address the issue that pre-defined gestures do not completely reflect user intent, we evaluated earlier work on user-defined gestures. As a result of these studies, we were also motivated to consider whether the addition of user-defined gesture interaction will reduce sign language learners' weariness and boost their motivation for learning sign language.

Therefore, the primary goal of our research is to investigate if user-defined gesture interaction affects ASL learning. In order to give users an immersive experience, we developed a VR-based system. To stimulate users' curiosity and boost their motivation, we also included a simple game with a user-defined hand gesture function. In addition, because there is a lack of user research on the subject, we used a questionnaire to survey users to investigate whether customised gesture interactions can actually inspire more people to learn sign language. Our main objective was to critically assess our system and gather user feedback on how their interaction with our system affected their learning experience.

3 System Components

This section provides an overview of the key components of the proposed system. The system's recommended workflow is shown in Fig. 1. When the user enters their account information through the login interface, the system initialises their location to the Instruction interface. After users familiarize themselves with the user introduction information of the Instruction interface, they study the ASL dictionary for five minutes and then visit the sign language game interface to play the game and increase their understanding of sign language through it.

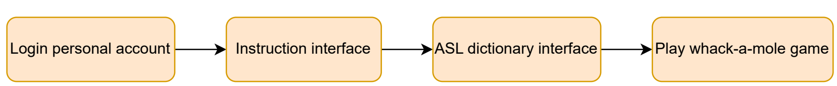


Fig. 1. The workflow of ASL learning system.

3.1 Learning Environment

The learning tools and the game for learning the numbers 0 to 9 in ASL are displayed in Fig. 2, with the user's viewpoint tilted 45 degrees to the left. The entire scene was created in Unity (2020.3.32f1). Regarding the user's engagement with the system, we used the eye-tracking functionality of the HTC Vive Pro and enable clicking or picking an object after 3 seconds of the user's fixed attention. An inbuilt camera that was connected to the PC via openCV (version 3.4.2) [4] was used to acquire the images. Regarding gesture detection and recognition, Mediapipe is used to detect the user's hand and extract a series of 21 points matching corresponding landmarks on the detected hand. The feature vector from this sequence is then supplied as input to the classifier, which is an MLP with three fully connected layers, implemented in Tensorflow 2.6.0 [8] and Python 3.6 [19]. We used an RTX3080 GPU on a standard PC to train the classifier. The study's objectives were satisfied with an overall recognition accuracy rate of over 90%, which is expected to offer a generally positive user experience.

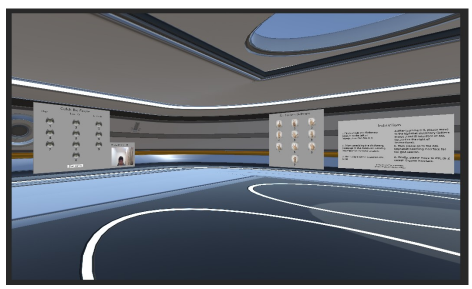


Fig. 2. The implemented ASL learning environment. **From left to right:** the Whack-a-Mole game; the ASL dictionary for the 0-9 digits; the Instructions interface.

3.2 Whack-a-Mole Game and User-defined Interface

We adopted the Whack-a-Mole game and implemented a sign language-based version of it, aiming to make learners more interested in the material, increasing their motivation and, eventually, their engagement with the process. In our game, as shown in Fig. 3(a), each location is marked by a unique numeric identifier. If the user signs correctly the current position of the gopher, one point is added; otherwise, no point is awarded. The total duration of the game is 30 seconds.

The user-defined interface is a feature that is hidden from the user while they are playing the game. It will only appear when a mole is killed and trigger a hitherto hidden functionality, calling for the user to specify a wake-up gesture to be utilized later in the game. In the example shown in Fig. 3(b), the user is given the special game skill to "Clear all Moles". At the end of one iteration of the course, the system will collect user-defined gesture data for 5 seconds, retrain the

(a) (b)

Fig. 3. (a) The Whack-a-Mole game for ASL learning; (b) the user defined interface.

recognition model, and the user-defined interface will be hidden again. Now that the user has picked their special skills, they may start playing the game again by clicking "Start" on the game screen, and they can use their newly acquired special game skill. The workflow of the user-defined hand gesture interface is shown in Fig. 4. Notice that the special mole sequence must match the last digit of the remaining play time for the user interface to be activated. In addition, the user must be able to recognise the special mole's number at a specific moment. When both requirements are satisfied, the user-defined interface will be displayed and the user will be awarded their special game skill.

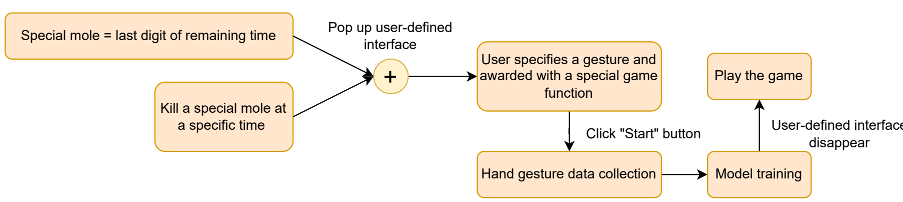


Fig. 4. The workflow of the user-defined hand gesture interface.

4 Experiments

To evaluate the system design, we adopted the user survey scheme proposed by Schrepp *et al.* [21], which comprises six evaluation factors: **Attractiveness**, **Efficiency**, **Perspicuity**, **Dependability**, **Stimulation**, **Novelty**. Each factor is further divided into six or seven more specialised issues. Table 1 displays the specific issues associated with each factor. Based on the users' scores on a scale of 1.00 to 5.00 on particular issues, we assessed the merits of the system in each factor.

We invited 15 users (M = 8, F = 7; aged between 19 and 21) to engage with our system, aiming at gathering user feedback to serve as the study's data source.

The majority of the users had either very limited or no prior understanding of ASL, or any other sign language. They were instructed to explore the system, adhering to the instructions in order to learn ASL in three stages: learning signs from the dictionary interface; improving comprehension at the learning interface; and assessing their learning three times on the game interface. As it can be challenging for some beginners to pick up so many motions quickly, users were merely required to learn the ASL 0-9 numerals.

Table 1. System evaluation questionnaire.

Attractiveness	Efficiency
Do users like or dislike the product? annoying / enjoyable good / bad unlikable / pleasing unpleasant / pleasant attractive / unattractive afriendly / unfriendly	Is it easy to understand how to use the product? Is it easy to get familiar with the product? not understandable / understandable easy to learn / difficult to learn complicated / easy clear / confusing
Perspicuity	Dependability
Is it possible to use the product fast and efficient? Does the user interface looks organized? fast / slow inefficient / efficient impractical / practical organized / cluttered	Does the user feel in control of the interaction? Is the interaction with the product secure and predictable? unpredictable / predictable obstructive / supportive secure / not secure meets expectations / does not meet expectations
Stimulation	Novelty
Is it interesting and exciting to use the product? Does the user feel motivated to further use the product? valuable / inferior boring / exiting not interesting / interesting motivating / demotivating	Is the design of the product innovative and creative? Does the product grab user's attention? creative / dull inventive / conventional usual / leading edge conservative / innovative

5 Result Analysis

The user evaluation is summarized in Fig. 5, the box-plots showing the Minimum, First Quartile, Median, Third Quartile, and Maximum, while the Mean is shown by an 'x'. The score distribution reflects generally positive feedback on the evaluation factors, all of which received mean scores greater than 3.00, while the overall system achieved a satisfactory average score of 3.42 (SD = 0.88) over the six factors. It is also interesting to note that some low scores (< 2.50) were given in all factors, the possible causes of which are discussed below.

Attractiveness: as shown in Fig. 6(a), the average score over the 7 questions is 3.16 (SD = 0.90). Each question has some scores lower than 2.50, possibly reflecting some lack of interaction with the users. For example, some users reported that animations should show up when the hand gesture was recognised correctly. Warnings should also be shown if no hand was detected, or when the hands were too close to the camera. Future improvement plans include the addition of more interactive features, such as moving backgrounds, scene changes, and animations.

Efficiency: as shown in Fig. 6(b), the average score is 3.40 (SD = 0.93). There were 3 users who gave a score of 5.00 on some questions. However, nearly one-third of them gave scores lower than 2.50 on each question, indicating that there

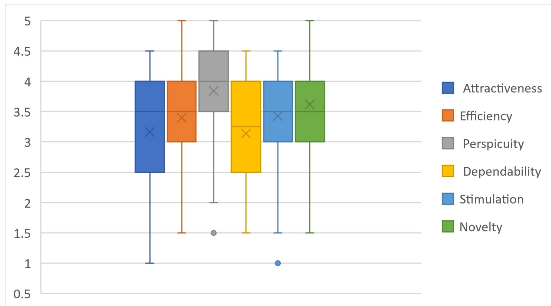


Fig. 5. Box-plots of the score values for each of the user survey's factors.

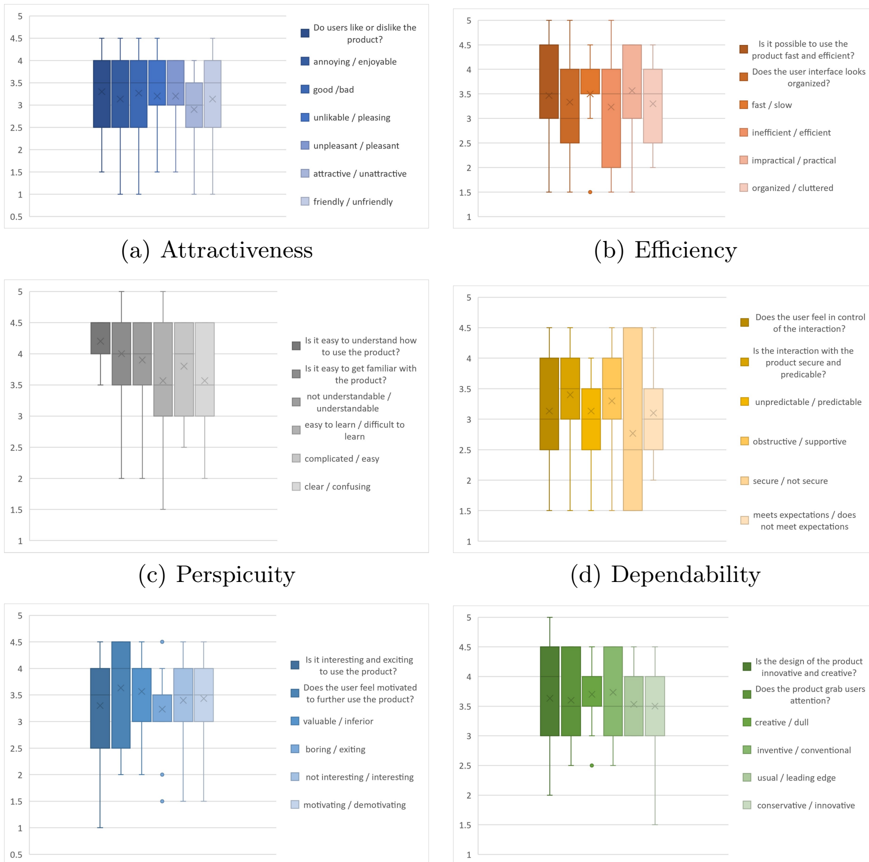
is still room for increasing the system's efficiency. According to user feedback, the practice interface was not so convenient to use, as the users had to click the button "Start" to check for correctness. Improvements could be made to automate this process, thus requiring less activity of this type from the users.

Perspicuity: as shown in Fig. 6(c), the average score of 3.84 (SD = 0.77) is the highest among the six factors, indicating that most users saw the system as easy to use, perhaps because of the simple design of the interface, which made it easy to use. Nevertheless, one user complained about the scene navigation system, overlooking apparently the navigation instructions button of the main menu. Thus, when this particular user tried to move to the ASL 0-9 dictionary, they did not know how to do it until we told them to look at the button. In a future study, the users will first be taught how to navigate the scene, before going into the main study of gestures.

Dependability: as shown in Fig. 6(d), the average score is 3.14 (SD = 0.90). Although most users gave scores higher than 3.50, a small number of users gave scores of 1.50. Perhaps this was because, in the practice interface, some users thought that they did the correct gesture but were judged as being wrong. The reason behind this can be performance issues of the gesture recognition model in the back-end, perhaps because the training data for some gestures might have been of poor quality, thus leading to low recognition accuracy rates for these gestures. Future improvements will aim at training recognition models that will be able to better generalize to natural physiological differences in users' hands.

Stimulation: as shown in Fig. 6(e), the average score for this factor is 3.43 (SD = 0.85), while the mean value of every question under this factor is above 3.25. It is also worthwhile to note that the first quartiles of all questions start from just below 3.00, and most of the users gave scores between 3.50 and 4.00. Nevertheless, were also some low scores in all questions, showing that there is still room for improvement. In particular, the interaction of the system could be designed more creatively, aiming at better inspiring the users.

Novelty: as shown in Fig. 6(f), the average score of this factor is 3.62 (SD = 0.70), suggesting a broadly positive reception, with the mean score in all questions at 3.50 or above. The second and the fourth questions have relatively



(a) Attractiveness (b) Efficiency

(c) Perspicuity (d) Dependability

(e) Stimulation (f) Novelty

Fig. 6. Box-plots of the scores for each subdivision of the six factors.

tight distributions, with interquartile range between 3.00 and 4.50. Overall, the results on this factor suggest that most of the users regarded the system as being innovative, with only a few of them perceiving it differently.

Fig. 7 shows that users typically had poor game scores at their initial attempt, with the exception of one outlier with a score of 18, who had the good fortune to activate the user-defined interface and obtain the hidden game skill on their first attempt. We note that the average, lowest, and highest user score all gradually increased at the second and third attempt, showing that the user's sign language proficiency increased. Additionally, several users claimed that this hidden feature might stimulate their interest in the game, and, implicitly, help them advance their sign language skills.

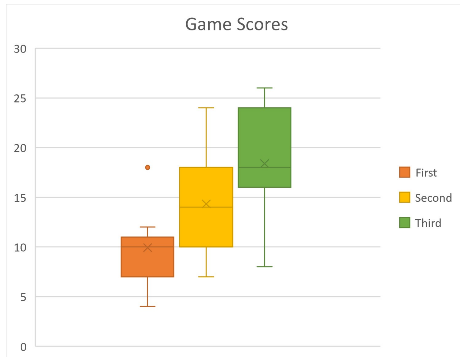


Fig. 7. Box-plots of the game scores for each of the three attempts.

6 Discussion

Using VR technology, we developed an immersive environment for learning ASL. We looked into whether a user-defined way of interaction could boost users' motivation to learn sign language. We evaluated this issue using the survey method proposed by Schrepp *et al.* [21].

For the user evaluation of the system we used six assessment factors. The survey's results on these six metrics demonstrate how well the user-defined interface for ASL learning operates and that it can genuinely satisfy user needs. Besides, because user-defined interactions are more experiential, most users seem to accept them, according to the analysis of environmental experience. However, the system still needs to be further optimised and adjusted for some functions to improve user experience, even though the majority of users are happy with user-defined interactions.

6.1 Limitations

Our work still has a number of limitations in terms of system design and implementation, as was already mentioned. Here, we summarise them for each assessment factor separately. **Attractiveness**: when sign language was correctly interpreted, some users complained that there weren't any animated clues. Animations in suggestions could enhance user experience. **Efficiency**: because users have to manually touch the start button to play the sign language game, some users complained that the user interface wasn't sufficiently automatic. **Perspicuity**: some users have complained that they were unsure of how to navigate the scenario. **Dependability**: while playing the game, some users claimed to have used the correct sign language, but the algorithm determined that they had not, giving them a lesser score. **Stimulation**: a small percentage of users gave low scores on the stimulation factor study, which suggests that there is still room for our system to be designed more creatively. **Novelty**: a small proportion of users felt that the system wasn't inventive enough, possibly because they thought the learning model was too simple.

On the other hand, our study has a number of methodological drawbacks. The user study and included only 15 participants. The invited people were between the ages of 19 and 21; there is no research on users in other age groups. To further test our methodology, we intend to enlist more individuals in future studies, who should come from a wider range of backgrounds (e.g., age). Ultimately, there is no specific evaluation indicator in the questionnaire survey on sign language acquisition. A uniform and standardised questionnaire-based assessment of sign language acquisition is needed for the next research.

7 Conclusion

A virtual environment that allows users to learn ASL through the use of user-defined hand gestures has been developed by our team. The user interface that is embedded in the virtual environment made it possible to most of the users to readily comprehend the workflow of the system, as well as each stage of the ASL learning process. The results of a user questionnaire that we carried out (N = 15) revealed that participants were, in general, pleased with the digital ASL learning system that we developed. In conclusion, the overall results provide credence to our original hypothesis, which stated that an increase in users' motivation to learn can be attributed to the usage of user-defined interaction modalities.

In the future, we will include in the system more interactive components, such as backdrop movement, scene changes, and animation prompts. In order to reduce the amount of human involvement required to control the system, we will also add more automatic settings. To help the user understand how to manipulate the objects in the scene, a follow-through user interface will be developed. In addition, a stronger gesture recognition model will be developed, enabling the inclusion of more sophisticated sign language instruction materials.

References

1. Adamo-Villani, N., Carpenter, E., Arns, L.: An immersive virtual environment for learning sign language mathematics. In: ACM SIGGRAPH 2006 educators program, pp. 20–es (2006)
2. Battistoni, P., Di Gregorio, M., Sebillo, M., Vitiello, G.: Ai at the edge for sign language learning support. In: IEEE HCC Conference. pp. 16–23 (2019)
3. Bheda, V., Radpour, D.: Using deep convolutional networks for gesture recognition in american sign language. arXiv preprint arXiv:1710.06836 (2017)
4. Bradski, G., Kaehler, A.: Opencv. DDJ of software tools **3**, 120 (2000)
5. Bragg, D., Caselli, N., Gallagher, J.W., Goldberg, M., Oka, C.J., Thies, W.: Asl sea battle: Gamifying sign language data collection. In: Proceedings of the 2021 CHI conference on human factors in computing systems. pp. 1–13 (2021)
6. Camgoz, N.C., Koller, O., Hadfield, S., Bowden, R.: Sign language transformers: Joint end-to-end sign language recognition and translation. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. pp. 10023–10033 (2020)

7. Dai, Q., Li, X., Geng, W., Jin, W., Liang, X.: Capg-myo: A muscle-computer interface supporting user-defined gesture recognition. In: Proceedings of the 9th ICCCM. pp. 52–58 (2021)
8. Dillon, J.V., Langmore, I., Tran, D., Brevdo, E., Vasudevan, S., Moore, D., Patton, B., Alemi, A., Hoffman, M., Saurous, R.A.: Tensorflow distributions. arXiv preprint arXiv:1711.10604 (2017)
9. Empe, N.A.A., Echon, R.C.L., Vega, H.D.A., Paterno, P.L.C., Jamis, M.N., Yabut, E.R.: Simbowika: A mobile and web application to learn filipino sign language for deaf students in elementary schools. In: (R10-HTC). pp. 1–6. IEEE (2020)
10. Estrada-Cota, I., Carreño-León, M.A., Sandoval-Bringas, J.A., Leyva-Carrillo, A.A., Quiroz, H.X.C.: Design of a web tool for teaching–learning of states and capitals of méxico through the mexican sign language. In: International Conference on Inclusive Technology and Education (CONTIE). pp. 174–179. IEEE (2021)
11. Goswami, T., Javaji, S.R.: Cnn model for american sign language recognition. In: ICCCE 2020, pp. 55–61. Springer (2021)
12. Jiang, X., Hu, B., Chandra Satapathy, S., Wang, S.H., Zhang, Y.D.: Fingerspelling identification for chinese sign language via alexnet-based transfer learning and adam optimizer. Scientific Programming (2020)
13. Kim, S., Ji, Y., Lee, K.B.: An effective sign language learning with object detection based roi segmentation. In: 2018 Second IEEE IRC. pp. 330–333. IEEE (2018)
14. Koh, J.I., Cherian, J., Taele, P., Hammond, T.: Developing a hand gesture recognition system for mapping symbolic hand gestures to analogous emojis in computer-mediated communication. ACM TiiS **9**(1), 1–35 (2019)
15. Pallavi, P., Sarvamangala, D.: Recognition of sign language using deep neural network. IJARCS **12**, 92–97 (2021)
16. Patricks, A.: Developing an accessible learning application for sign language (c) (2022)
17. Phan, H.D., Ellis, K., Dorin, A., Olivier, P.: Feedback strategies for embodied agents to enhance sign language vocabulary learning. In: Proceedings of the 20th ACM International Conference on Intelligent Virtual Agents. pp. 1–8 (2020)
18. Piumsomboon, T., Clark, A., Billinghurst, M., Cockburn, A.: User-defined gestures for augmented reality. In: IFIP Conference on HCI. pp. 282–299. Springer (2013)
19. Python, W.: Python. Python Releases for Windows **24** (2021)
20. Schioppo, J., Meyer, Z., Fabiano, D., Canavan, S.: Sign language recognition: Learning american sign language in a virtual environment. In: Extended Abstracts of the CHI Conference on Human Factors in Computing Systems. pp. 1–6 (2019)
21. Schrepp, M., Hinderks, A., Thomaschewski, J.: Applying the user experience questionnaire (ueq) in different evaluation scenarios. In: International Conference of Design, User Experience, and Usability. pp. 383–392. Springer (2014)
22. Takayama, Y., Ichikawa, Y., Shizuki, B., Kawaguchi, I., Takahashi, S.: A user-based mid-air hand gesture set for spreadsheets. In: Asian CHI Symposium. pp. 122–128 (2021)
23. Wu, H., Wang, Y., Liu, J., Qiu, J., Zhang, X.L.: User-defined gesture interaction for in-vehicle information systems. Multimed. Tools Appl. **79**(1), 263–288 (2020)
24. Wu, H., Wang, Y., Qiu, J., Liu, J., Zhang, X.: User-defined gesture interaction for immersive vr shopping applications. BIT **38**(7), 726–741 (2019)
25. Zhang, F., Bazarevsky, V., Vakunov, A., Tkachenka, A., Sung, G., Chang, C.L., Grundmann, M.: Mediapipe hands: On-device real-time hand tracking. arXiv preprint arXiv:2006.10214 (2020)