Developing and Evaluating a Novel Gamified Virtual Learning Environment for ASL

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. The use of sign language is a highly effective way of communicating with individuals who experience hearing loss. Despite extensive research, many learners find traditional methods of learning sign language, such as web-based question-answer methods, to be unengaging. This has led to the development of new techniques, such as the use of virtual reality (VR) and gamification, which have shown promising results. In this paper, we describe a gamified immersive American Sign Language (ASL) learning environment that uses the latest VR technology to gradually guide learners from numeric to alphabetic ASL. Our hypothesis is that such an environment would be more engaging than traditional web-based methods. An initial user study showed that our system scored highly in some aspects, especially the hedonic factor of novelty. However, there is room for improvement, particularly in the pragmatic factor of dependability. Overall, our findings suggest that the use of VR and gamification can significantly improve engagement in ASL learning.

Keywords: Human Computer Interaction · ASL Learning · VR

1 Introduction

Sign language is a visual language that uses hand gestures and facial expressions to convey meaning. It is primarily used for communication with individuals who are deaf or hard of hearing or who experience difficulty speaking. Learning sign language is important for several reasons. Firstly, it enables better communication and social interaction with the hearing-loss community, thereby promoting inclusion and understanding. By learning sign language, one can break down communication barriers and establish meaningful connections with individuals who might otherwise feel excluded. Secondly, learning sign language has been shown to have numerous cognitive benefits, including enhancing cognitive development and language skills [4]. It is widely acknowledged that learning a second language has cognitive benefits, and the same is true for sign language. Finally, for individuals who experience hearing or speech impairments, sign language can serve as a crucial mode of communication, allowing them to participate more fully in society. Despite the importance of learning sign language, traditional web-based methods of learning have not been able to generate much interest among learners, partly because of a lack of novelty. Therefore, there is a need for more engaging and innovative approaches to learning sign language that can increase user engagement and promote effective learning.

To improve the user experience of ASL learning, we developed a VR-based learning environment that incorporated a Whack-a-Mole type of game, inspired by the ASL game Sea Battle used by Bragg *et al.* for data collection [6]. We then conducted a user study, utilising a questionnaire proposed by Schrepp *et al.* [17], to evaluate the user experience of our system. To the best of our knowledge, there have been no previous user studies focused on the user experience of ASL learning from numeric to alphabetic in a gamified VR environment. Hence, our main research question was: ***"Were users satisfied with the ASL learning experience from numeric to alphabetic in a gamified VR environment?"***. By conducting this user study, we aimed to gain insight into how users experienced our system and identify areas where improvements could be made. Ultimately, we hoped to demonstrate that incorporating gamification and VR technology into ASL learning can enhance user satisfaction and engagement. Our main contributions are as follows:

1. We successfully created an immersive virtual environment that supports ASL learning from numeric to alphabetic, incorporating a Whack-a-Mole type of game. Our system provides a unique and engaging approach to ASL learning, which we believe can enhance user satisfaction and engagement.
2. Our user study provided initial evidence that our approach has the potential to improve some aspects of user experience. These findings indicate that incorporating immersive elements and games into ASL education may be a promising direction for improving user satisfaction and learning outcomes.

2 Related Work

Sign language recognition: The recognition of sign language through deep learning and computer vision has been studied by various researchers. Bheda *et al.* [3] proposed a method that uses deep convolutional neural networks to recognize ASL gestures. Kim *et al.* [12] presented a novel approach that employs an object detection network for the region of interest (ROI) segmentation to preprocess input data for sign language recognition. Battistoni *et al.* [2] described a method that allows for monitoring the learning progress of ASL alphabet recognition through CNNs. Jiang *et al.* [11] proposed a transfer learning-based approach for identifying fingerspelling in Chinese Sign Language. Camgoz *et al.* [7] introduced a transformer-based architecture that jointly learns Continuous Sign Language Recognition and Translation. Zhang *et al.* [20] proposed a real-time on-device hand tracking pipeline called MediaPipe Hands for AR/VR applications. Goswami *et al.* [10] created a new dataset and trained a CNN-based model for recognizing hand gestures in ASL. Finally, Pallavi *et al.* [13] developed a deep learning model based on the YOLOv3 architecture, reporting high recognition rates for the ASL alphabet. These studies demonstrate the potential of deep learning and computer vision techniques in improving accessibility for individuals with hearing impairments.

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.

Sign language applications: The article discusses various research studies related to sign language applications. Bantupalli *et al.* [1] created a vision-based system to translate sign language into text to improve communication between signers and non-signers. Schnepp *et al.* [16] developed an animated sign language dictionary for caregivers to learn communication with residents who use sign language. Samonte [15] created an e-tutor system to assist instructors in teaching sign language. Economou *et al.* [9] designed a Serious Game to help adults learn sign language and bridge the communication gap between hearing-impaired and able-hearing people. Wang *et al.* [19] designed a sign language game with user-defined features and found that gamified sign language learning can improve the user's learning experience. These studies suggest that dictionary searches and gamification can improve the learning experience, and influenced the design choices for our system.

We developed a virtual reality system that offers an immersive and interactive learning experience for sign language. To improve the user experience, we incorporated a quiz and a small game into the system. Given the dearth of research in this area, we conducted user interviews using a questionnaire to evaluate users' satisfaction with ASL learning from numeric to alphabetic in the system. Our objectives were twofold: to thoroughly evaluate the performance of our system and to investigate users' experiences with it.

3 User Interface of VR Environment

This section provides an overview of the main components of our user interface (UI) and highlights the main features of our VR environment. The UI is comprised of four different modules designed to facilitate effective ASL learning.

1. The **Instructions** module, which consists of six basic steps, provides users with an overview of the ASL learning process and guides them through the initial stages of the programme.
2. The **Sign Language Dictionaries** module, which enables users to consult and search for the signs of numbers or letters. This module serves as a reference tool for users as they progress through the learning process.
3. The **Quiz** module, which contains question-answer quizzes that allow users to test their signing skills and self-assess their level of competence. This

module serves as a valuable feedback mechanism for users and encourages them to actively engage with the learning material.
4. The **Whack-a-Mole Game** module, which is to increase user motivation and engagement with the learning process. This module presents users with a fun and interactive way to practice their ASL skills, reinforcing their learning and providing a welcome break from more traditional learning methods.

Together, these four modules work in concert to provide users with a comprehensive and engaging VR-based ASL learning experience. By incorporating elements of gamification and interactivity into our VR environment, we hope to improve user satisfaction and facilitate more effective ASL learning outcomes.

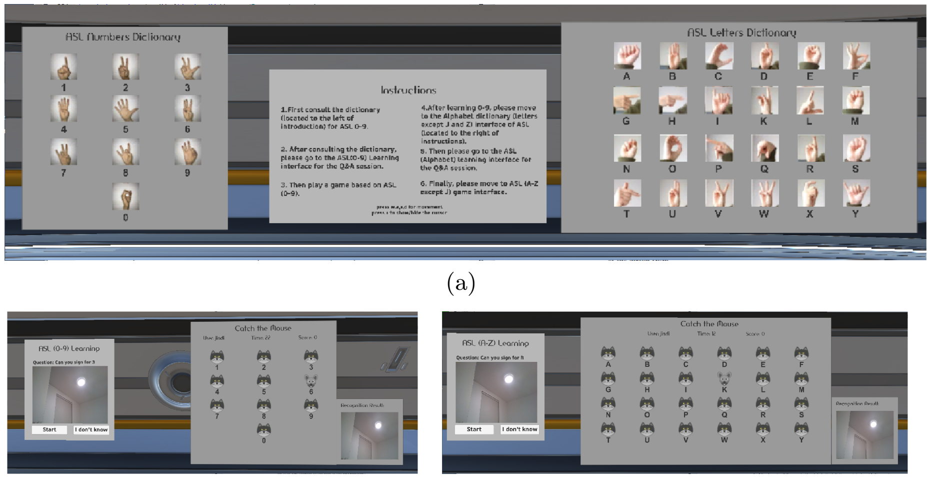
We separated the scene of the immersive environment into two parts. Adopting the concept of a simple to complex learning process, the first part is for learning the numeric ASL, something that is considered a relatively easy task. The second part of the scene is for the more challenging task of learning the alphabetic ASL, excluding J and Z, which require dynamic gesturing.

Fig. 1(a) shows the initial view of the user when entering the VR environment, which includes the **Instructions** and **Sign Language Dictionary** interfaces. Fig. 1(b) shows the **Quiz** and **Whack-a-Mole Game** interfaces of numerical ASL learning, which are located to the left of the numerical ASL dictionary. Fig. 1(c) shows the **Quiz** and **Whack-a-Mole Game** interfaces of alphabetic ASL learning, which are located to the right of the alphabetic ASL dictionary.

The scene was developed in Unity 2020.3.32f1, and user interaction was facilitated through eye tracking using HTC Vive Pro. After 3 seconds of fixed attention, users can click or select objects in the scene. An integrated camera was used to acquire images; openCV (version 3.4.2) [5] was used for image processing on a PC. Hand gestures were detected using Mediapipe, which extracted a feature vector of 21 points corresponding to landmarks on the detected hand. An MLP consisting of 3 fully connected layers was implemented in Python 3.6 [14] and Tensorflow 2.6.0 [8] for gesture recognition. The classifier was trained on a standard PC with an RTX3080 GPU, achieving recognition accuracy rates above 90%, deemed sufficient to ensure a smooth user experience in our study.

4 User Study Design

In order to evaluate the immersive environment design, we adopted the user survey scheme proposed by Schrepp *et al.* [17], which is commonly used to evaluate user experience in human-computer interaction systems. It consists of six evaluation factors, called *scales*: **Attractiveness**, **Efficiency**, **Perspicuity**, **Dependability**, **Stimulation**, **Novelty**. Each scale is further divided into four or six *items*, as shown in Table 1. We evaluated the proposed VR environment, on all scales and items, on a 7-point Likert scale ranging from -3 (fully agree with a negative term) to +3 (fully agree with a positive term), and studied the user feedback against the benchmark proposed in [18]. In that paper, the authors analysed a large database of questionnaire responses and derived the benchmark intervals shown in Table 2. These intervals correspond to the distribution:



(a)

(b) (c)

Fig. 1. The implemented immersive virtual environment. (a) **Left:** the numeric ASL sign language dictionary. **Centre:** Instructions interface. **Right:** the A-Y except for J sign language dictionary. (b) The numeric ASL learning quiz (left) and game (right). (c) The alphabetic ASL learning quiz (left) and game (right).

Table 1. Summary of the user experience questionnaire.

Attractiveness	**Perspicuity**
A1: annoying / enjoyable	**P1**: not understandable / understandable
A2: good / bad	**P2**: easy to learn / difficult to learn
A3: unlikable / pleasing	**P3**: complicated / easy
A4: unpleasant / pleasant	**P4**: clear / confusing
A5: attractive / unattractive	
A6: friendly / unfriendly	
Efficiency	**Dependability**
E1: fast / slow	**D1**: unpredictable / predictable
E2: inefficient / efficient	**D2**: obstructive / supportive
E3: impractical / practical	**D3**: secure / not secure
E4: organized / cluttered	**D4**: meets expectations / does not meet expectations
Stimulation	**Novelty**
S1: valuable / inferior	**N1**: creative / dull
S2: boring / exciting	**N2**: inventive / conventional
S3: not interesting / interesting	**N3**: usual / leading edge
S4: motivating / demotivating	**N4**: conservative / innovative

- **Excellent**: In the range of the 10% best results.
- **Good**: 10% of results better, 75% of results worse.
- **Above average**: 25% of results better, 50% of results worse.
- **Below average**: 50% of results better, 25% of results worse.
- **Bad**: In the range of the 25% worst results.

We conducted the user study obtaining feedback from 15 participants, 8 males and 7 females, aged between 19 and 21 years old, who had little or no prior experience with ASL or any other sign language. At the start of the session,

Table 2. Benchmark intervals for the user experience scales.

	Attractiveness	Perspicuity	Efficiency	Dependability	Stimulation	Novelty
Excellent	≥ 1.75	≥ 1.78	≥ 1.90	≥ 1.65	≥ 1.55	≥ 1.40
Good	[1.52, 1.75)	[1.47, 1.78)	[1.56, 1.90)	[1.48, 1.65)	[1.31, 1.55)	[1.05, 1.40)
Above average	[1.17, 1.52)	[0.98, 1.47)	[1.08, 1.56)	[1.14, 1.48)	[0.99, 1.31)	[0.71, 1.05)
Below average	[0.70, 1.17)	[0.54, 0.98)	[0.64, 1.08)	[0.78, 1.14)	[0.50, 0.99)	[0.30, 0.71)
Bad	<0.70	<0.54	<0.64	<0.78	<0.50	<0.30

participants had the freedom to explore the system and consult the Instructions module. Then, each participant followed a six stages learning process:

1. Learn numeric ASL for 3 minutes from corresponding dictionary module.
2. Improve numeric ASL comprehension for 3 minutes in numeric quiz module.
3. 30 seconds on numeric ASL game module.
4. Learn alphabetic ASL from corresponding dictionary module for 3 minutes.
5. Improve alphabetic ASL literacy for 3 minutes in alphabetic quiz module.
6. 30 seconds on alphabetic ASL game module.

5 Result Analysis

Fig. 2 shows the average scores for the six scales, denoted by 'x', plotted over a colour code of the corresponding benchmark interval. For each scale, the minimum and the maximum of the average scores on its individual items are also shown. In Fig. 3, the box plots show the minimum, first quartile, median, third quartile, and maximum, for each individual item of each scale.

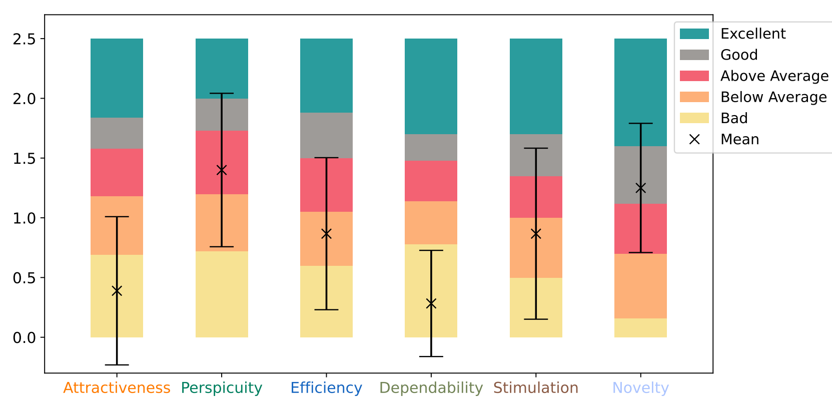
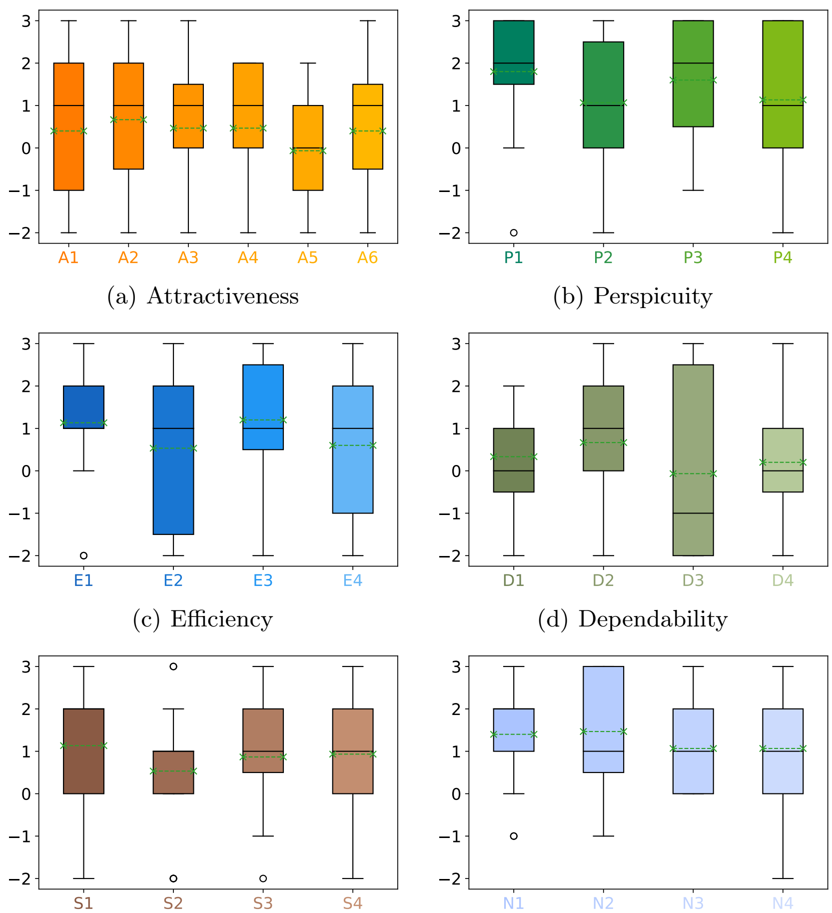


Fig. 2. Benchmark intervals for the six scales

Attractiveness: The mean value of the user scores is 0.39 (SD = 1.24), placing it in the "Bad" category, indicating that their overall impression of the VR



(a) Attractiveness (b) Perspicuity
(c) Efficiency (d) Dependability
(e) Stimulation (f) Novelty

Fig. 3. Box-plots of the scores for each item of the six scales.

environment was not favourable, and the system requires further improvements. Notably, the average score for item **A5**, shown in Fig. 3(a), is slightly below 0, which suggests that the users did not find the system particularly appealing. This may be because the learning environment relies on 2D user interfaces, whereas incorporating 3D elements may be more visually engaging for users. Therefore, we plan to integrate 3D user interfaces in future iterations of the ASL learning environment, aiming at enhancing its attractiveness.

Perspicuity: The average score is 1.40 (SD = 1.28), placing it in the "Above average" category, indicating that users perceive the VR environment as clear and understandable, facilitating their ASL learning experience. However, it seems that some of the users may have encountered some problems when using the environment, possibly due to their unfamiliarity with VR devices, and they may require some initial training.

Efficiency: In the "Below average" category, the average score is 0.87 (SD = 1.27). We note that, while the average score over the whole scale is slightly below average, analysis of individual item scores shows that our VR environment adequately fulfills some users' requirements. In particular, users found the system easy to use (as reflected by item **E1**) and believed that they could practice ASL effectively in the scenario (as reflected by item **E3**), see Fig. 3(c).

Dependability: In the "Bad" category, the average score is 0.28 (SD = 0.89). That means that the VR environment's dependability needs significant improvement. Despite the low overall average score, some users still believed that on individual items, particularly **D2** and **D4**, the system adequately fulfilled their requirements, see Fig. 3(d).

Stimulation: In the "Below average" category with an average score of 0.87 (SD = 1.43). Even though the score is slightly lower than average, the large variance indicates that some users find the learning environment stimulating. As shown in Fig. 3(e), the first quartile of all items is non-negative, indicating that a majority of users have a consistently favourable outlook regarding this scale.

Novelty: In the "Good" category with an average value of 1.25 (SD = 1.08). Again, the first quartile of all items is non-negative, see Fig. 3(f), indicating a consistently favourable view from a majority of users. They perceive the VR environment as a novel and innovative way of learning ASL.

6 Conclusion

We have developed a VR system for learning numeric and alphabetic ASL and conducted a questionnaire-based user study to evaluate the user experience of learning ASL in the system. We found that to some extent it satisfied some user satisfaction factors, however, the system needs further development to enhance user experience, especially on the factors of attractiveness and dependability.

There are several limitations to our ASL learning system, which have been discussed for each scale of user experience separately. The identified shortcomings include a lack of animated hints; an interface that requires users to actively press a start button to commence an action; difficulty in moving around the VR scene; a relatively large number of incorrect judgments of correct signs, i.e., many false negatives; user expectations for a more creatively designed system; and an overall perception that the learning task was too easy. Additionally, the user study included 15 only participants, primarily between the ages of 19 and 21, and there was a complete lack of research on users in other age groups.

To address these limitations, we plan to revise the content, design, and implementation of the system as follows: add more interactive elements; implement automatic settings; create a follow-through user interface; develop a more robust sign recognition model; and include more sophisticated sign language learning material. We also plan to recruit a larger and more diverse group of participants for a follow-up user study.

References

1. Bantupalli, K., Xie, Y.: American sign language recognition using deep learning and computer vision. In: 2018 IEEE International Conference on Big Data (Big Data). pp. 4896–4899. IEEE (2018)
2. Battistoni, P., Di Gregorio, M., Sebillo, M., Vitiello, G.: Ai at the edge for sign language learning support. In: 2019 IEEE International Conference on Humanized Computing and Communication (HCC). pp. 16–23. IEEE (2019)
3. Bheda, V., Radpour, D.: Using deep convolutional networks for gesture recognition in american sign language. arXiv preprint arXiv:1710.06836 (2017)
4. Bialystok, E., et al.: Bilingualism in development: Language, literacy, and cognition. Cambridge University Press (2001)
5. Bradski, G., Kaehler, A.: Opencv. Dr. Dobb's journal of software tools **3**, 120 (2000)
6. 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)
7. 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)
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. Economou, D., Russi, M.G., Doumanis, I., Mentzelopoulos, M., Bouki, V., Ferguson, J.: Using serious games for learning british sign language combining video, enhanced interactivity, and VR technology. Journal of Universal Computer Science **26**(8), 996–1016 (2020)
10. Goswami, T., Javaji, S.R.: Cnn model for american sign language recognition. In: ICCCE 2020, pp. 55–61. Springer (2021)
11. 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** (2020)
12. Kim, S., Ji, Y., Lee, K.B.: An effective sign language learning with object detection based roi segmentation. In: 2018 Second IEEE International Conference on Robotic Computing (IRC). pp. 330–333. IEEE (2018)
13. Pallavi, P., Sarvamangala, D.: Recognition of sign language using deep neural network. International Journal of Advanced Research in Computer Science **12**, 92–97 (2021)
14. Python, W.: Python. Python Releases for Windows **24** (2021)
15. Samonte, M.J.C.: An assistive technology using FSL, speech recognition, gamification and online handwritten character recognition in learning statistics for students with hearing and speech impairment. In: Proceedings of the 2020 The 6th International Conference on Frontiers of Educational Technologies. pp. 92–97 (2020)
16. Schnepp, J., Wolfe, R., Brionez, G., Baowidan, S., Johnson, R., McDonald, J.: Human-centered design for a sign language learning application. In: Proceedings of the 13th ACM International Conference on PErvasive Technologies Related to Assistive Environments. pp. 1–5 (2020)
17. 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)

18. Schrepp, M., Thomaschewski, J., Hinderks, A.: Construction of a benchmark for the user experience questionnaire (UEQ). International Journal of Interactive Multimedia and Artificial Intelligence **4**(4), 40–44 (2017)
19. Wang, J., Ivrissimtzis, I., Li, Z., Zhou, Y., Shi, L.: User-defined hand gesture interface to improve user experience of learning american sign language. In: International Conference on Intelligent Tutoring Systems. pp. 479–490. Springer (2023)
20. 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)