



OPEN ACCESS

EDITED BY

Athina Papadopoulou,
New York Institute of Technology, United States

REVIEWED BY

Bencie Woll,
University College London, United Kingdom
Maria Tagarelli De Monte,
Università degli Studi Internazionali di Roma,
Italy
Hamed Mohammad Hosseini,
Islamic Azad University Central Tehran
Branch, Iran

*CORRESPONDENCE

Santiago Berrezueta-Guzman,
✉ s.berrezueta@tum.de

RECEIVED 09 May 2025

ACCEPTED 28 August 2025

PUBLISHED 22 September 2025

CITATION

Berrezueta-Guzman S, Daya R and Wagner S
(2025) Virtual reality in sign language education:
opportunities, challenges, and the road ahead.
Front. Virtual Real. 6:1625910.
doi: 10.3389/frvir.2025.1625910

COPYRIGHT

© 2025 Berrezueta-Guzman, Daya and Wagner.
This is an open-access article distributed under
the terms of the [Creative Commons Attribution
License \(CC BY\)](#). The use, distribution or
reproduction in other forums is permitted,
provided the original author(s) and the
copyright owner(s) are credited and that the
original publication in this journal is cited, in
accordance with accepted academic practice.
No use, distribution or reproduction is
permitted which does not comply with these
terms.

Virtual reality in sign language education: opportunities, challenges, and the road ahead

Santiago Berrezueta-Guzman*, Refia Daya and Stefan Wagner

Chair of Software Engineering, Technical University of Munich, Munich, Germany

Sign language (SL) is an essential mode of communication for Deaf and Hard-of-Hearing (DHH) individuals. Its education remains limited by the lack of qualified instructors, insufficient early exposure, and the inadequacy of traditional teaching methods. Recent advances in Virtual Reality (VR) and Artificial Intelligence (AI) offer promising new approaches to enhance sign language learning through immersive, interactive, and feedback-rich environments. This paper presents a systematic review of 55 peer-reviewed studies on VR-based sign language education, identifying and analyzing five core thematic areas: (1) gesture recognition and real-time feedback mechanisms; (2) interactive VR environments for communicative practice; (3) gamification for immersive and motivating learning experiences; (4) personalized and adaptive learning systems; and (5) accessibility and inclusivity for diverse DHH learners. The results reveal that AI-driven gesture recognition systems integrated with VR can provide real-time feedback, significantly improving learner engagement and performance. However, the analysis highlights critical challenges: hardware limitations, inconsistent accuracy in gesture recognition, and a lack of inclusive and adaptive design. This review contributes a comprehensive synthesis of technological and pedagogical innovations in the field, outlining current limitations and proposing actionable recommendations for developers and researchers. By bridging technical advancement with inclusive pedagogy, this review lays the foundation for next-generation VR systems that are equitable, effective, and accessible for sign language learners worldwide.

KEYWORDS

Virtual Reality, sign language education, gesture recognition, deaf and hard-of-hearing (DHH), AI in learning, immersive learning environments, accessibility and inclusivity

1 Introduction

Sign languages (SLs) are natural human languages with their own phonology, morphology, syntax, and pragmatics, born and evolving within Deaf communities around the world. Although they make use of the visual-spatial modality rather than the auditory-vocal channel, sign languages are linguistic systems just as rich and expressive as spoken languages, and in no way are they simple gestural codes or mere translations of the surrounding spoken tongue (Stokoe, 2005; Klima and Bellugi, 1979). Each SL, whether American Sign Language (ASL), British Sign Language (BSL), or Brazilian Sign Language (LIBRAS), to name a few, developed independently, with unique grammatical structures and regional variation, entirely independent of the ambient spoken language.

As with any natural language, fluency in a sign language grants its users full access to education, employment, social services, and cultural life. Yet, despite official recognition in many countries, SL education remains unevenly available: globally, only about 5%–10% of

Deaf and Hard-of-Hearing (DHH) children acquire a sign language from birth, while the vast majority encounter their first structured exposure to SL much later, often through under-resourced school programs or self-directed learning (Newport, 1988; Humphries et al., 2012). These delays and limitations in early access have profound effects on cognitive and linguistic development, on educational attainment, and on social inclusion. Addressing these inequities requires both expanding access to qualified instructors and exploring innovative pedagogical approaches—including the use of immersive and interactive technologies such as Virtual Reality (VR) and Artificial Intelligence (AI)—to bring high-quality, accessible SL learning to all who need it.

A significant portion of the DHH population learns SL later in life due to a lack of early exposure, the scarcity of qualified instructors, and the limited integration of SLs into formal education systems (Rho et al. (2020); Novaliendry et al. (2023)). Traditional teaching approaches, such as textbooks, video lessons, and instructor-led classrooms, often fail to capture SL's visual-spatial and expressive complexities, limiting learners' engagement, comprehension, and retention (Wen et al. (2024); Shaw et al. (2023); Wang et al. (2024b)).

Recent advancements in immersive technologies, particularly the integration of Virtual Reality (VR) and Artificial Intelligence (AI), offer new opportunities for transforming SL education (Wang et al. (2024a)). VR enables three-dimensional, interactive, and context-rich environments that simulate real-life communication scenarios, provide visual demonstrations from multiple perspectives, and support real-time feedback (Alam et al. (2024); Parmaxi (2020); Bisio et al. (2023)). When enhanced with AI-powered gesture recognition, these systems can analyze learner performance, deliver personalized feedback, and create gamified, adaptive experiences that support diverse learners (Bansal et al. (2021); Sagayam and Hemanth (2017)). Despite this promise, current implementations remain fragmented, with notable limitations in scalability, accessibility, cultural inclusivity, and pedagogical depth (Rho et al. (2020); Jiang et al. (2019); Sabbella et al. (2024)).

This paper presents a systematic review of VR-based SL education. It synthesizes findings across 55 peer-reviewed studies to identify the key applications, technological innovations, and pedagogical strategies employed in the field. This review focuses on five thematic areas: (1) gesture recognition and feedback mechanisms; (2) interactive VR environments for practice; (3) immersive learning through gamification; (4) personalized learning and adaptive systems; and (5) accessibility and inclusivity.

By mapping these areas against the current challenges and future needs of learners and educators, this paper informs the design of next-generation learning tools that are effective, ethical, and equitable. Additionally, it provides practical recommendations for developers, educators, and researchers and proposes directions for future work to support scalable, inclusive, and culturally responsive SL learning through immersive technologies.

2 Background

The intersection of SL education and emerging immersive technologies has gained increasing attention as educators and

researchers seek innovative ways to address persistent challenges in accessibility, pedagogy, and learner engagement.

2.1 Overview of sign language education

Sign languages are acquired through distinct developmental trajectories depending on learners' age, hearing status, and learning context. Deaf children born to signing parents (approx. 5%–10% of the population) acquire a sign language natively, following the same milestones as hearing children do in spoken language acquisition (Schick et al., 2010). Early exposure—whether from Deaf family members, early intervention programs, or signing daycare environments—ensures robust linguistic and cognitive development, and has become increasingly available through universal newborn hearing screening and family-centred early intervention services (Lillo-Martin and Henner, 2021).

For deaf children of hearing parents, structured SL instruction often begins in infancy or toddlerhood via early intervention specialists and specialized preschool programs, rather than waiting for the first school classroom. Contemporary models integrate Deaf mentors, parent coaching, and bilingual-bicultural curricula, recognizing that home signing systems (“home signs”) lack the full grammatical complexity of natural SLs but can scaffold early communication until richer input is provided (Branson and Miller, 2005; Morford and Hänel-Faulhaber, 2011).

By contrast, hearing adult learners—including parents, interpreters, and professionals—typically begin structured SL study well after childhood. Adult courses range from university credit classes and community workshops to online modules and immersion camps. Pedagogical approaches for adults emphasize explicit instruction in SL grammar (handshape, location, movement, non-manual markers), cultural norms, and functional fluency, often combining video-based modeling, in-person practice with Deaf signers, and increasingly, digital platforms with interactive feedback (Quinto-Pozos, 2011). Research shows that adult learners benefit from multimodal input and scaffolded practice, but progress varies widely based on intensity, motivation, and access to native models.

Across all learner groups, effective SL education requires

- *Rich linguistic input* from fluent signers (for both child and adult learners);
- *Immediate, targeted feedback* on phonological parameters and facial grammar;
- *Culturally responsive pedagogy* that situates signs within Deaf community norms and Deaf-hearing intercultural communication.

Despite expanded early intervention and adult instruction options, gaps remain in qualified Deaf instructor availability, equitable program funding, and evidence-based curriculum standards—challenges that immersive technologies like VR/AI are uniquely poised to address.

For the deaf community, SL is a fundamental means of communication. Despite its importance, learning and teaching SL remain challenging due to limited course availability, a shortage of qualified instructors, and the language's inherent complexity. For

example, American Sign Language (ASL) is a visual-gestural language with unique linguistic structures. However, unlike most spoken languages, only about 5%–10% of deaf individuals acquire ASL from birth in households where SL is naturally used. The majority are instead introduced to it later in life, often through formal education programs rather than through natural exposure in their home environments [Newport \(1988\)](#).

Although SLs are recognized as official languages in many countries, the levels of fluency are low because of the barriers to acquisition and the absence of practical approaches to teaching [Rho et al. \(2020\)](#). Most students learn SL for the first time at school and not at home, and, historically, SL was banned from the classroom, with deaf children having to teach themselves from other deaf children [Novaliendry et al. \(2023\)](#).

SL education often relies on classroom teaching, textbooks, and video-based resources. While classroom instruction can provide personalized feedback, it is frequently limited by a lack of qualified instructors and difficulty offering consistent one-on-one support [Wang et al. \(2024b\)](#). Textbooks and videos are more flexible and accessible but lack interactivity and fail to represent the spatial and expressive features of SL, such as hand orientation, movement, and facial expressions, that are essential for accurate communication [Wen et al. \(2024\)](#); [Shaw et al. \(2023\)](#). Research indicates that while some learners are fortunate to be taught by native signers, many are instructed by non-native users, which introduces variability in teaching effectiveness [Newport \(1988\)](#).

Approximately 90%–95% of deaf children are born to hearing parents who do not use SL, resulting in limited or no exposure to a fully developed linguistic system during critical developmental years. While some children develop ‘home signs’—gestural systems created within families to facilitate essential communication—these systems lack the grammatical complexity of natural languages and do not support complete language acquisition. This early lack of linguistic input can lead to cognitive and language development delays, making subsequent mastery of a formal SL significantly more difficult [Newport \(1988\)](#).

Moreover, feedback is critical in SL acquisition. Despite its importance, there is currently no standardized approach to integrating feedback into SL learning methodologies [Shaw et al. \(2023\)](#); [Senthamarai \(2018\)](#); [Novaliendry et al. \(2023\)](#). Learners are typically required to self-monitor and self-correct, which is particularly difficult in the absence of visual or interactive cues [Alam et al. \(2024\)](#); [Hore et al. \(2017\)](#); [Cui et al. \(2017\)](#); [Huang et al. \(2018\)](#). While some technological solutions aim to address this gap, current SL recognition systems still face challenges in accurately interpreting dynamic gestures and facial expressions—both essential components of meaning in SL communication [Wen et al. \(2024\)](#).

2.2 Virtual reality in sign language education

Initially developed for entertainment, Virtual Reality (VR) has increasingly gained traction in educational contexts and is now applied across various domains. Its growing relevance in education stems from its ability to simplify complex concepts and extend

learning opportunities beyond geographical boundaries [Alam et al. \(2024\)](#).

In language learning, VR supports the development of communicative competence, cultural awareness, critical thinking, and kinesthetic understanding, particularly beneficial when learning spatially grounded languages such as SL [Parmaxi \(2020\)](#). VR also facilitates real-time interaction and task-based learning, allowing learners to practice in realistic scenarios and perform authentic communicative tasks [Chen et al. \(2025\)](#).

Furthermore, studies show that VR enhances motivation and learner autonomy, with serious games and gamified experiences contributing to higher engagement and persistence compared to traditional methods such as lectures or video-based instruction [Ortiz et al. \(2017\)](#); [Suh et al. \(2016\)](#); [Damianova and Berrezueta-Guzman \(2025\)](#).

Aligned with experiential learning theory, VR fosters active learning by immersing students in problem-solving and decision-making tasks, which supports memory retention, deeper comprehension, and the development of cognitive and metacognitive skills [Rho et al. \(2020\)](#); [Parmaxi \(2020\)](#). This is particularly valuable in SL education, where learners must grasp spatial relationships, precise hand movements, and visual cues. Traditional methods often lack interactivity and fail to represent the dynamic, three-dimensional nature of SL [Wang et al. \(2024a\)](#); [Novaliendry et al. \(2023\)](#).

A core advantage of VR in this context is its integration with motion capture technology, which enables real-time gesture recognition and immediate feedback on accuracy. Devices such as Leap Motion controllers and Oculus hand tracking enhance the ability of VR applications to track hand positions and analyze sign execution [Bisio et al. \(2023\)](#); [Rho et al. \(2020\)](#). These tools ensure learners receive precise visual feedback, helping them correct errors and internalize accurate signing techniques.

Machine learning models, including Convolutional Neural Networks (CNNs) and Hidden Markov Models (HMMs), are employed to interpret gestures and evaluate performance in real-time. These models allow learners to practice signs in VR environments where errors are detected automatically and corrective feedback is provided instantly. These algorithms assess hand postures and inform the user whether the gesture was performed correctly, encouraging repetition and reinforcement [Alam et al. \(2024\)](#); [Bansal et al. \(2021\)](#).

To enhance realism and instructional support, some VR systems integrate signing avatars—driven by motion-capture or advanced pose-estimation tools such as AlphaPose, Azure Kinect, and MediaPipe—to demonstrate correct sign execution and prompt real-time user adjustments ([Alam et al., 2024](#)). For example, Bansal et al.’s CopyCat uses these vision-only pipelines alongside Hidden Markov Models trained on keypoint data to verify ASL sentence production in real time, achieving over 90% word-level accuracy without specialized gloves ([Bansal et al., 2021](#)). Such AI-powered avatar feedback loops substantially improve comprehension, retention, and practical SL application in immersive learning environments.

Some applications often simulate everyday conversations and social scenarios, bridging the gap between theoretical instruction and real-life communication [Novaliendry et al. \(2023\)](#). Such contextual learning not only reinforces vocabulary but also supports the development of fluency and situational understanding [Wang et al. \(2024b\)](#).

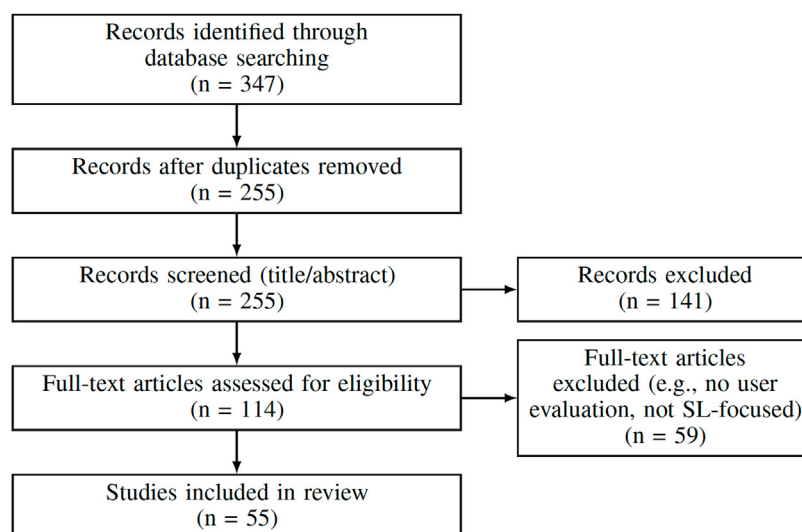


FIGURE 1
PRISMA flow diagram illustrating the study selection process.

Engagement in VR-based SL education is further enhanced through gamification and adaptive learning strategies. Studies have shown that game-like environments increase learner motivation, persistence, and time spent in learning compared to traditional formats Wang et al. (2024a); Shohieb (2019); Wang et al. (2024c). AI-driven avatars within these environments often serve as interactive tutors, offering immediate performance feedback and visual scoring systems to support learner understanding Bisio et al. (2023); Shi et al. (2014).

Although VR for SL education is still evolving, ongoing developments suggest a promising future that includes haptic feedback to reinforce correct handshapes and multi-modal recognition systems capable of interpreting facial expressions—an essential component of SL grammar and meaning Alam et al. (2024); Wen et al. (2024).

3 Methodology

This study employed a systematic literature review to examine current applications of Virtual Reality (VR) in Sign Language (SL) education. The methodology involved a structured search strategy, selection criteria, and categorization to identify key themes and trends across recent works.

3.1 Literature search strategy

Our search included electronic databases such as IEEE Xplore, ACM Digital Library, Scopus, Web of Science, and Google Scholar. The search terms used were combinations of keywords such as “sign language learning,” “virtual reality,” “VR-based education,” “gesture recognition,” “deaf education,” and “interactive environments.” Boolean operators (AND, OR) were applied to refine the search.

The inclusion criteria were:

- Studies published between 2010 and 2025.
- Peer-reviewed journal articles, conference proceedings, or book chapters.
- Studies that include technical or pedagogical evaluations of VR systems in SL learning.

Exclusion criteria included:

- Studies not written in English.
- Works that did not involve VR or SL learning directly.
- Purely theoretical or opinion papers without implementation or user evaluation.

Figure 1 provides a PRISMA flowchart which shows that 55 relevant studies were identified for full-text analysis, and Figure 2 illustrates how these studies are distributed by year and type of language.

3.2 Categorization of literature

Each article was reviewed for its objectives, methodology, technology, and findings. Five major categories emerged from this analysis:

1. *Gesture Recognition and Feedback Mechanisms*: Studies focusing on AI/ML-based recognition systems for sign detection and real-time feedback.
2. *Interactive VR Environments for Practice*: Applications that simulate conversational or task-based scenarios for SL use.
3. *Immersive Learning Through Gamification*: Studies implementing game elements to increase engagement and retention.
4. *Personalized Learning and Adaptive Systems*: Systems offering user-specific content progression, replay features, or real-time adaptation.
5. *Accessibility and Inclusivity*: Works addressing the needs of DHH users through avatars and translation tools.

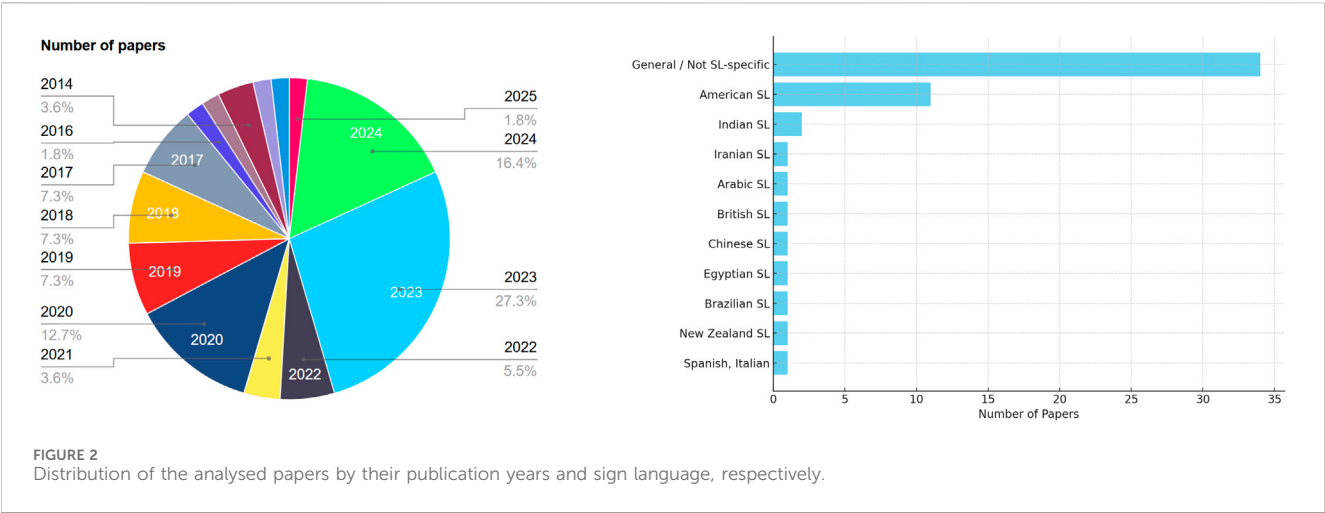


TABLE 1 Summary of key findings and limitations of virtual reality in sign language learning.

Category	Principal Findings	Limitations
Gesture recognition and feedback mechanisms	Advanced gesture recognition using CNNs, LSTMs, HMMs, and 3D CNNs enables real-time sign evaluation with high accuracy; visual feedback facilitates learning; Unity and Oculus are commonly used	Hardware dependency, limited gesture datasets, no haptic feedback, challenges with dynamic gestures, and low facial expression tracking accuracy.
Interactive VR environments for practice	Gamified VR with contextual scenarios enables immersive practice; avatar-based multiplayer design fosters realistic interactions and engagement	Mostly individual-use; limited multiplayer integration, and lacks culturally adapted communication models.
Immersive learning through gamification	Gamification improves motivation, engagement, and retention; gesture-triggered puzzles and animations reinforce sign-memory links	Some gamified elements are superficial; they may not address learner autonomy, confusion in 3D environments, or match novice users' needs.
Personalized learning and adaptive systems	Some systems allow manual replay, camera control, and non-linear navigation; conceptual frameworks for adaptive learning exist	Lack of AI-based adaptation, minimal real-time performance tracking, and no intelligent feedback loops.
Accessibility and inclusivity	VR avatars and visual feedback support DHH learners; they report improved motivation, and the proposed multilingual environments	High hardware cost, limited support for interpreter avatars, absence of standardized academic signs, and limited support for users with multiple disabilities.

4 Key applications of virtual reality in sign language learning

The final categorization enabled a structured synthesis of the findings and facilitated the construction of Table 1 that summarizes principal applications, limitations, and key studies.

4.1 Gesture recognition and feedback mechanisms

AI-driven gesture recognition in VR-based SL learning focuses on accurately detecting hand shapes, movements, and spatial configurations. Convolutional Neural Networks (CNNs) are commonly used for static gesture recognition, while Long Short-Term Memory (LSTM) networks are more effective for modeling temporal dynamics Goyal (2023); Elhagry and Elrayes (2021); Wu et al. (2018); Zhu et al. (2017). Hidden Markov Models (HMMs) have also shown strong performance,

particularly in structured, real-time educational settings involving young learners, with one system achieving over 90% user-independent accuracy and outperforming transformer-based models Alam et al. (2024); Bansal et al. (2021). More advanced models, including 3D CNNs, deformable 3D CNNs, and convolutional LSTMs, have been applied to classify gestures in video and multimodal streams, achieving high accuracy—some above 98 %—and enabling real-time feedback with minimal latency Zhang and Wang (2019); Zhang et al. (2020); Ur Rehman et al. (2022); Qi et al. (2024); Sabbella et al. (2024).

The tracking method used in most cases relies on vision-based input, which ranges from basic webcam feeds Zhang and Wang (2019) to structured video datasets Liu et al. (2023) and MediaPipe Holistic landmark detection Goyal (2023); Monisha et al. (2023). The systems operating with hardware devices make use of the Oculus Quest 2 or Meta Quest 2 Tju and Shalih (2024); Carnevale et al. (2022); Ramos-cosi et al. (2023) to record hand position, orientation, and grip parameters for immediate evaluation

and Electromyography (EMG) signal acquisition to detect muscle-based gestures Wu et al. (2018). The analysis of gesture detection systems also includes evaluations for hand skeletonization, posture recognition, and dual-hand interaction design Qi et al. (2024); Sagayam and Hemanth (2017); Jiang et al. (2019). Most research relies on data-driven modeling, yet specific conceptual frameworks stress user-centered evaluation and gesture datasets standardization Cortes et al. (2024); Sabbella et al. (2024); Gupta et al. (2020).

The majority of gesture recognition systems included real-time feedback mechanisms, which allowed users to correct their mistakes and enhance their learning outcomes Qi et al. (2024); Sagayam and Hemanth (2017); Ur Rehman et al. (2022). The applications show gesture recognition outcomes through on-screen visual cues or VR immersive interfaces, improving user interaction and engagement Wu et al. (2018); Sabbella et al. (2024).

Several Unity-based implementations applied gesture models to 3D spaces that included characters and animations together with gameplay elements Qi et al. (2024); Sabbella et al. (2024). These systems functioned as gamified VR teaching tools for children, using Oculus tracking and Unity rendering for SL instruction Tju and Shalih (2024); Carnevale et al. (2022); Ramos-cosi et al. (2023).

Most systems use visual feedback, but they do not include haptic feedback, like vibrations or physical sensations, to guide users. Even though some systems use gloves or VR controllers, these devices are not used to give touch-based responses. Adding this kind of physical feedback could improve learning, especially for activities that rely on movement and touch Tju and Shalih (2024); Carnevale et al. (2022).

Despite strong performance in controlled environments, most systems showed limitations related to hardware reliance, data collection variability, and sensitivity to environmental conditions Tju and Shalih (2024); Jiang et al. (2019); Wu et al. (2018). These factors indicate that there is still a requirement for scalable and adaptive models that perform well with users from different backgrounds in various environments and languages Sagayam and Hemanth (2017); Sabbella et al. (2024).

4.2 Interactive VR environments for practice

Virtual reality environments with immersive features enable SL students to practice through role-playing, simulation, and gamified interaction Alam et al. (2024); Mazhari et al. (2022). Learners can enhance their sign accuracy by practicing gestures in meaningful contexts because these environments provide experiential cues and task completion feedback Rho et al. (2020). Multiple VR-based systems for SL education use game mechanics and scenario progression to boost user engagement and motivation Shaw et al. (2023); Quandt et al. (2020). The user's progress, task success, and virtual object interaction respond directly to gesture input in these environments, thus establishing a natural feedback system between learning and doing Economou et al. (2020); Adamo-Villani et al. (2006).

VR applications in SL now integrate multiplayer features that enable users to engage in real-time interaction, cooperative activities, and dialogue-based practice Berrezueta-Guzman and Wagner (2025). The systems recognize SL as a social communication method, establishing virtual environments that mimic natural communication scenarios. Research involving DHH participants shows that multiplayer platforms enhance collaborative learning

while enabling non-verbal communication and peer engagement through visual methods such as gestures and avatar interactions Luna (2023). However, in several cases, SL was used alongside or replaced by text chat or speech-to-text tools, especially when users did not share the same SL Luna et al. (2023).

The broader educational VR environments have shown through multiplayer design that this approach leads to better immersion and collaboration and improved contextual learning results Darejeh (2023). The direct implementation of multiplayer SL interaction within VR learning systems would create more authentic communication experiences, which would help students develop fluent and responsive communication skills and social self-assurance Maroukas et al. (2023).

4.3 Immersive learning through gamification

Several serious games for SL education demonstrate that gameplay elements, such as animations, scores, or avatar reactions, enhance memory retention by linking gestures to specific outcomes Shaw et al. (2023); Alam et al. (2024). This immediate feedback reinforces motor memory and comprehension, making learning more effective and engaging. Studies show that when signs trigger in-game actions or puzzles, learners perceive the experience as more meaningful and enduring than traditional methods Mazhari et al. (2022); Economou et al. (2020).

VR systems have been developed where American, British, and Iranian SLs are practiced through mini-games, object interaction, and goal-oriented puzzles Shaw et al. (2023); Alam et al. (2024); Mazhari et al. (2022); Economou et al. (2020). Some environments are designed to teach domain-specific vocabulary, such as mathematical signs, by embedding them in contextual fantasy scenarios Adamo-Villani et al. (2006). Others provide real-time feedback on gesture shape, timing, and execution, helping users internalize corrections through repeated gameplay Bisio et al. (2023); Quandt et al. (2020). These implementations demonstrate how serious games can facilitate SL practice while maintaining learner engagement through meaningful task design, immediate reinforcement, and immersive interaction.

4.4 Personalized learning and adaptive systems

Some approaches use environmental control and user-directed navigation to simulate personalized learning. For example, VR-based platforms for learning Brazilian Sign Language (LIBRAS) provided users with the ability to control camera angles, slow down animations, and replay signs at their own pace, which helped with visual self-pacing and independent practice without AI-driven adaptation Ferreira Brega et al. (2014). Similarly, systems incorporating learning style frameworks, like Kolb's experiential model, provide learners with multiple content formats (visual, auditory, reading) and let them choose non-linear learning paths based on preference Horváth (2020). These approaches support learner autonomy but are based on manual customization rather than intelligent adjustment based on performance.

More advanced approaches propose the use of machine learning techniques for real-time adaptation. For instance, one AR/VR

framework uses classification and clustering algorithms to track learner progress and recommend appropriate modules or feedback strategies on the fly. However, this system was not sign-language-specific [Al-Ansi et al. \(2023\)](#). Nonetheless, there remains a gap between gesture-based feedback and truly adaptive, performance-driven learning pathways across the literature reviewed.

Current systems tend to focus more on autonomy rather than on intelligent adaptation. Few applications incorporate real-time analysis of gesture accuracy, fluency, or error trends to adjust lesson difficulty or suggest targeted practice. This gap indicates a clear opportunity for future work: integrating AI-driven user modeling and performance-based progression could significantly enhance the effectiveness of VR SL instruction by aligning content more closely with each learner's evolving needs.

4.5 Accessibility and inclusivity

The visual aspects of VR environments match the communication requirements of DHH users by providing them with immersive learning experiences. Various systems show how gesture-based interaction combined with avatar-signed content enables more intuitive learning environments. A VR-based SL learning tool implemented motion-captured avatars alongside physical interaction through pinch gloves to let deaf children communicate naturally [Adamo \(2007\)](#).

Several studies highlight user demand for accessible VR tools tailored to DHH learners. Survey-based research showed that 64% of deaf university students expressed interest in using VR for lectures, with 80% emphasizing the need for virtual SL interpreters. The same participants demonstrated better motivation and understanding when they engaged in virtual reality learning scenarios [Gunarhadi et al. \(2024\)](#).

Another study explored the development of multilingual sign translation in VR chatrooms, aiming to support more inclusive real-time communication. However, it remained at the prototype stage without empirical validation [Teófilo et al. \(2018\)](#).

5 Challenges

5.1 Hardware and software limitations

VR-based SL learning systems experience limitations due to hardware and software restrictions. The literature shows an ongoing problem in finding scalable and affordable input and output devices that deliver high-fidelity performance [Sagayam and Hemanth \(2017\)](#); [Jiang et al. \(2019\)](#). Vision-based systems, while more accessible than sensor-laden systems, experience tracking problems and occlusion events along with latency issues and performance degradation when lighting conditions are poor [Sabbella et al. \(2024\)](#); [Monisha et al. \(2023\)](#). Systems utilizing Leap Motion, webcams, or standalone VR headsets like Oculus Quest 2 report issues with limited tracking range, spatial resolution, and precision, particularly when recognizing detailed hand configurations or subtle finger movements [Carnevale et al. \(2022\)](#); [Gupta et al. \(2020\)](#).

The accuracy-enhancing hardware solutions, including wearables, EMG sensors, and specialized gloves, bring their difficulties [Zhang and Wang \(2019\)](#). Most gesture recognition

methods demand complex hardware configurations while creating physical restrictions on users and generating variable signal quality, reducing system practicality, restricted scalability, minimal system interoperability, and restricted availability consistency [Wu et al. \(2018\)](#); [Ramos-cosi et al. \(2023\)](#).

The software side faces additional restrictions because computational limitations make implementing advanced gesture recognition models in real-time systems difficult. The processing capabilities of standalone VR headsets prove insufficient to execute high-complexity models that use CNN-LSTM hybrids or transformer-based architectures [Alam et al. \(2024\)](#); [Huang et al. \(2018\)](#).

Additionally, several studies highlight the lack of infrastructure and technical support in educational settings, which prevents the deployment of VR systems for widespread classroom use [Bisio et al. \(2023\)](#); [Cortes et al. \(2024\)](#). The operation and maintenance of VR tools demand specialized training for educators, which creates additional hurdles for adoption [Parmaxi \(2020\)](#). Creating high-quality 3D signing avatars presents developmental challenges that require extensive resources and time while also complicating expanding support for multiple SLs and user groups [Quandt et al. \(2020\)](#).

Finally, despite the promise of web-based and AI-enhanced solutions, many VR applications still overlook fundamental accessibility needs, particularly for DHH users with intersecting impairments [Creed et al. \(2023\)](#). Inclusive hardware-software design that addresses the needs of diverse learners remains largely underexplored.

5.2 Accuracy of gesture recognition

VR-based SL learning systems face accurate and consistent gesture recognition. Deep learning and motion tracking have produced significant progress, but many applications still face challenges when interpreting gestures in real-world immersive situations [Goyal \(2023\)](#); [Liu et al. \(2023\)](#); [Hore et al. \(2017\)](#).

The fundamental problem stems from the challenging task of identifying static versus dynamic signs. Identifying dynamic gestures becomes challenging because of movement variations, hidden gestures, and hand motion overlap. The problem worsens because the SL community lacks sufficient standardized datasets that show regional and stylistic variations in signs as researched in Indian and Egyptian SL studies [Elhagry and Elrayes \(2021\)](#); [Bisio et al. \(2023\)](#). The restricted nature of these systems to static vocabularies prevents them from supporting the complete range of continuous SL communication [Zhang and Wang \(2019\)](#); [Rho et al. \(2020\)](#).

Vision-based systems show high sensitivity to lighting conditions, background elements, and hand position changes [Jiang et al. \(2019\)](#); [Zhu et al. \(2017\)](#). At the same time, commodity VR equipment from Oculus Quest and Leap Motion fails to detect precise finger movements and facial expressions [Sagayam and Hemanth \(2017\)](#); [Sabbella et al. \(2024\)](#); [Alam et al. \(2024\)](#).

Deep learning models, including CNNs, LSTMs, and transformers, need extensive, well-annotated datasets for effective generalization. Current datasets show two major limitations: they either contain small amounts of data or do not provide enough time-based information for continuous SL recognition, resulting in overfitting and performance issues [Zhang et al. \(2020\)](#); [Cui et al. \(2017\)](#). These constraints are especially limiting when systems

depend on manual segmentation or weak supervision, as seen in several gesture pipelines Qi et al. (2024); Rho et al. (2020); Huang et al. (2018). Some studies have even found VR-based gesture recognition to underperform compared to traditional 2D methods Shaw et al. (2023).

Finally, the lack of integrated real-time feedback mechanisms remains a significant barrier to effective learning. Learners who do not receive prompt corrective information risk maintaining incorrect gestures, which create barriers to learning and skill development retention Wen et al. (2024); Quandt et al. (2020).

The solution demands three main components: dataset development with high quality and inclusivity, neural architecture hybridization for spatial-temporal feature detection, and real-time feedback integration Bansal et al. (2021); Tju and Shalih (2024).

5.3 User experience and learning effectiveness

While VR can make learning more engaging, several studies show it does not constantly improve things like memory or long-term learning Wang et al. (2024a). Being immersed in a 3D, game-like environment may capture students' interest. Still, it can also be confusing and make learning harder when the instructions or tasks are too complex Wang et al. (2024b); Economou et al. (2020).

Usability and personalization are critical yet often underdeveloped aspects of VR-based SL learning environments. Many systems fail to follow user-centered or iterative development processes, resulting in interfaces that confuse learners while contradicting their needs Adamo (2007).

While interactive avatars and 3D elements can enhance visual clarity, their effectiveness hinges on high design standards, usability, and user acceptance Ferreira Brega et al. (2014); Quandt et al. (2020). Moreover, standard VR platforms often lack accommodation for individual learning styles, affective preferences, and prior knowledge, essential for learner satisfaction and retention Horváth (2020); Al-Ansi et al. (2023). Although personalized VR/AR tools are conceptually promising, their implementation demands sophisticated analytics and adaptive systems that remain largely absent from current applications.

The implementation of gamification approaches in education produces varying levels of effectiveness. Research shows that motivational and engagement levels increase through gamification, yet superficial elements only provide meaningful learning benefits when psychological needs of autonomy, competence, and relatedness are specifically addressed Ortiz et al. (2017); Suh et al. (2016). Contextual adaptation is critical; generic game elements may fail to support novice users or learners from diverse educational backgrounds Shi et al. (2014); Economou et al. (2020).

The existing research demonstrates insufficient development of multiplayer and collaborative experiences that specifically address deaf and hard-of-hearing (DHH) users Luna (2023). The needs of DHH participants are not addressed through standard communication strategies, group dynamics, and co-located interaction designs, which restrict their involvement in social VR and educational simulations. In remote learning contexts, tools to support hearing parents in learning SL for their deaf children are especially scarce Quandt et al. (2020).

The effectiveness of VR-based SL learning relies on more than technological innovation, even though VR provides promising new interaction methods.

5.4 Ethical and cultural considerations

A new ethical concern is that AI systems may be biased, especially when personalizing learning. Research shows that people can have different experiences with AI based on gender, which risks reinforcing unfair treatment Wang et al. (2024c). Since SL learning relies on personalized feedback, biased AI could harm trust, exclude minority groups, and discourage users from continuing.

The ethical dimension of VR identity representation is a new, unexamined area that needs attention. Users tend to interact with avatars that do not match their cultural background and personal identity, which may result in feelings of exclusion and limited self-expression. Research indicates that identity satisfaction varies based on avatar customization options, suggesting that inclusive avatar design is essential for fostering belonging and user acceptance in educational VR Ramos-cosi et al. (2023).

The lack of qualified interpreters presents an infrastructural and ethical challenge to providing equal access to knowledge and academic participation. VR-based interpretation systems and avatar-led instruction can fill this gap, but only if designed with ethical inclusivity and linguistic accuracy in mind Gunarhadi et al. (2024).

Undereducation and semi-literacy among deaf individuals stem from mainstream learning system exclusions and insufficient support for ongoing educational opportunities Shohieb (2019). SL remains inaccessible as a standard school subject while fluent teacher availability remains limited, thus creating ongoing barriers to learning, especially in regions with insufficient resources Novaliendry et al. (2023).

Ethical design means more than supporting inclusion, representation, and equal access. Developers need to understand the diverse needs of deaf communities, including differences in language, gender, culture, and education. This calls for inclusive design, involving users in the development process, and checking AI systems for bias. VR tools for SL learning will only reach their full potential if they are built with strong ethical standards.

6 Future directions

Several important areas still need more research and development to fully unlock the educational benefits of VR-based learning.

6.1 Advances in AI for gesture recognition

The success of VR-based sign language learning depends largely on how accurately and quickly systems can recognize gestures. Future research should focus on using advanced deep learning models, such as CNNs, LSTMs, Transformers, and attention mechanisms, to better capture the movements and timing involved in continuous signing. There is also a strong need for large, diverse datasets that reflect real-world variation, including regional differences, facial expressions, and

two-handed signs Goyal (2023); Elhagry and Elrayes (2021); Wu et al. (2018).

In addition, future systems should use multimodal input—combining hand motion with facial cues, voice, or muscle signals (EMG)—to improve recognition across different learning contexts Zhu et al. (2017); Alam et al. (2024). Improving the speed and efficiency of these models on standard VR headsets is key for real-time feedback. Collaboration between AI experts and sign language linguists will be essential to ensure these systems reflect the full richness and diversity of sign languages Bansal et al. (2021); Zhang et al. (2020); Zhang and Wang (2019); Qi et al. (2024); Huang et al. (2018).

6.2 Integration with Augmented Reality (AR)

Augmented Reality (AR) can enhance VR-based SL learning by enabling practice in real-world, context-rich environments. Using AR on mobile devices or smart glasses, learners can receive real-time feedback, visual overlays, and avatar demonstrations without needing a full VR setup. This makes learning more flexible and accessible, especially for mobile or blended learning approaches Al-Ansi et al. (2023); Alam et al. (2024).

For instance, learners could get immediate feedback while signing in everyday situations, such as ordering food or asking for directions, making applying skills in real life easier. AR could also help parents and teachers learn SL alongside deaf children in daily routines, promoting more inclusive and family-centered learning experiences Rho et al. (2020); Wen et al. (2024).

6.3 Open-source development and community initiatives

Many existing VR sign language systems were developed as standalone projects, limiting their reuse and scalability. To increase global impact, future work should focus on open-source platforms, modular toolkits, and standardized APIs for gesture recognition, avatars, and feedback. Publicly available datasets and reproducible models will support broader adoption, validation, and continuous improvement Quandt et al. (2020); Adamo (2007); Gunarhadi et al. (2024).

Involving Deaf and Hard-of-Hearing (DHH) users as co-creators ensures these tools meet real-world needs and reflect linguistic and cultural diversity. Future platforms should also allow user-generated content, allowing learners and educators to build and share custom lessons and learning paths Creed et al. (2023); Alam et al. (2024).

6.4 Ethical AI and representation in sign language VR

Ethical design must be central to developing VR SL learning tools to ensure fairness and inclusivity. Future systems should address bias in AI models through fairness audits, diverse training data, and precise feedback mechanisms. Avatar design should also reflect various identities, including different body types, skin tones, genders, and signing styles Ramos-cosi et al. (2023); Gunarhadi et al. (2024); Creed et al. (2023).

Developers should follow participatory design practices involving DHH users, educators, and interpreters at every stage. Clear ethical guidelines for data use, privacy, and accessibility are essential when designing for children or underserved communities. Future research should also prioritize culturally responsive content and expand support for SL, particularly in low-resource settings Shohieb (2019); Wang et al. (2024c); Novaliendry et al. (2023).

7 Recommendations for developers

Based on the analysis of recent research and system implementations in VR-based SL learning, we provide the following recommendations to developers aiming to build inclusive, effective, and scalable learning environments:

7.1 Prioritize inclusive and participatory design

- Involve DHH individuals throughout the design and evaluation process to ensure cultural relevance, accessibility, and usability.
- Allow for extensive avatar customization to support identity expression and increase user comfort and representation across age, ethnicity, gender, and disability.
- Ensure compatibility with assistive technologies (e.g., screen readers, closed captions) to support multi-disability access.
- Audit AI models for bias based on gender, language, or signing style to avoid reinforcing inequities.

7.2 Enhance gesture recognition with robust, multimodal AI

- Use hybrid deep learning models (e.g., CNN-LSTM, 3D CNNs with attention) capable of processing static and dynamic gestures.
- Train models on diverse datasets reflecting real-world variation in signing styles, regional dialects, and continuous sentence structures.

7.3 Design real-time feedback mechanisms

- Implement immediate visual or auditory cues to help learners self-correct gestures.
- Use gamified reinforcement, such as animated reactions or scoring systems, to increase motivation.
- Integrate haptic feedback (e.g., vibration) to support kinesthetic reinforcement of signs.

7.4 Support personalization and adaptive learning

- Incorporate AI-based learner modeling to adapt content difficulty and exercises based on performance.

- Offer differentiated feedback strategies tailored to learning styles or goals (e.g., expressive fluency vs. receptive accuracy).

7.5 Optimize for cost-effective and scalable hardware

- Design for low-cost, standalone headsets (e.g., Meta Quest 2) or mobile devices to reduce barriers to access.
- Ensure performance on mid-range processors by optimizing model size and latency.
- Minimize hardware dependency (e.g., gloves or EMG sensors) to ensure broader deployment.

7.6 Ensure open access and reusability

- Release gesture datasets, source code, and VR environments under open licenses to encourage collaboration and transparency.
- Document best practices and evaluation metrics to foster replicability and longitudinal study comparisons.

Implementing these recommendations will help developers close gaps in accessibility, learning effectiveness, and inclusivity and drive progress in VR-based sign language education.

8 Limitations of this review

While this review aimed to provide a comprehensive and structured synthesis of the current VR-based sign language education landscape, several methodological limitations must be acknowledged.

First, publication bias may have influenced the results, as studies reporting successful or innovative outcomes are more likely to be published than those with inconclusive or negative findings. This could have skewed the overall interpretation of effectiveness and feasibility in favor of more optimistic representations.

Second, the language scope was limited to English-language publications, excluding potentially relevant studies published in other languages. Given the global nature of sign language communities and technological development, this introduces a geographic and linguistic bias that may have overlooked valuable contributions from non-English-speaking regions.

These limitations suggest that the findings of this review should be interpreted with caution and considered part of an evolving body of research that would benefit from continued meta-analytical work, multilingual inclusion, and participatory analysis frameworks.

9 Conclusion

This paper has comprehensively reviewed the current state, opportunities, and challenges of using Virtual Reality (VR) for sign language (SL) education. We identified five key application areas through systematic analysis—gesture recognition, interactive

practice, gamified learning, personalized instruction, and accessibility—and assessed their strengths and limitations.

Our findings show that VR, especially when paired with AI-powered gesture recognition, offers promising solutions to key challenges in SL learning, such as the shortage of qualified instructors, low engagement, and limited feedback. Gamification and avatar-based interaction further support motivation and retention. However, several barriers remain. Hardware costs, limited gesture recognition accuracy, and the lack of adaptive learning features restrict scalability. Ethical and cultural concerns—including algorithmic bias and underrepresentation of diverse SLs—must also be addressed to ensure true inclusivity.

Future systems must focus on robust datasets, participatory and user-centered design, and AI-driven personalization to move the field forward. Building open-source, affordable, and culturally responsive tools is key to expanding access, especially in underserved regions.

Realizing the full potential of VR in SL education will require close collaboration among technologists, educators, linguists, and Deaf and Hard-of-Hearing (DHH) communities. With inclusive and thoughtful design, VR can become a powerful tool to enhance language learning, promote accessibility, and support learners worldwide.

Author contributions

SB-G: Writing – review and editing, Funding acquisition, Supervision. RD: Methodology, Investigation, Writing – original draft. SW: Writing – review and editing, Validation, Supervision.

Funding

The author(s) declare that financial support was received for the research and/or publication of this article. This research was financially supported by the TUM Campus Heilbronn Incentive Fund 2024 of the Technical University of Munich, TUM Campus Heilbronn. We gratefully acknowledge their support, which provided the essential resources and opportunities to conduct this study.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The author(s) declare that no Generative AI was used in the creation of this manuscript.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated

References

- Adamo, N. (2007). A virtual learning environment for deaf children: design and evaluation.
- Adamo-Villani, N., Carpenter, E., and Arns, L. L. (2006). 3d sign language mathematics in immersive environment.
- Al-Ansi, A. M., Jabooob, M., Garad, A., and Al-Ansi, A. (2023). Analyzing augmented reality (ar) and virtual reality (vr) recent development in education. *Soc. Sci. Humanit. Open* 8, 100532. doi:10.1016/j.ssaho.2023.100532
- Alam, M. S., Lamberton, J., Wang, J., Leannah, C., Miller, S., Palagano, J., et al. (2024). Asl champ!: a virtual reality game with deep-learning driven sign recognition. *Comput. and Educ. X Real.* 4, 100059. doi:10.1016/j.cexr.2024.100059
- Bansal, D., Ravi, P., So, M., Agrawal, P., Chadha, I., Murugappan, G., et al. (2021). "Copycat: using sign language recognition to help deaf children acquire language skills," in Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems, New York, NY, USA (New York, NY, USA: Association for Computing Machinery).
- Berrezueta-Guzman, S., and Wagner, S. (2025). Immersive multiplayer vr: unreal engine's strengths, limitations, and future prospects. *IEEE Access* 13, 85597–85612. doi:10.1109/access.2025.3570166
- Bisio, A., Yeguas-Bolivar, E., Aparicio Martínez, P., Redel, M., Pinzi, S., Rossi, S., et al. (2023). Training program on sign language: social inclusion through virtual reality in isense project. 104–109. doi:10.1109/MetroXRINE58569.2023.10405777
- [Dataset] Branson, J., and Miller, D. (2005). Damned for their difference: the cultural construction of deaf people as disabled. *Scand. J. Disabil. Res.* 7, 129–132. doi:10.1080/15017410510032244
- Carnevale, A., Mannocchi, I., Sassi, M., Carli, M., De Luca, G., Longo, U. G., et al. (2022). Virtual reality for shoulder rehabilitation: accuracy evaluation of oculus quest 2. *Sensors* 22, 5511. doi:10.3390/s22155511
- Chen, W., Berrezueta-Guzman, S., and Wagner, S. (2025). "Task-based role-playing vr game for supporting intellectual disability therapies," in 2025 IEEE International Conference on Artificial Intelligence and eXtended and Virtual Reality (AIxVR), Santiago Berrezueta-Guzman, 16 Dec 2024 (IEEE), 159–164.
- Cortes, D., Bermejo Gonzalez, B., and Juiz, C. (2024). The use of cnns in vr/ar/mr/xr: a systematic literature review. *Virtual Real.* 28, 154. doi:10.1007/s10055-024-01044-6
- Creed, C., Al Kalbani, M., Theil, A., Sarcara, S., and Williams, I. (2023). Inclusive ar/vr: accessibility barriers for immersive technologies. *Univers. Access Inf. Soc.* 23, 1–15. doi:10.1007/s10209-023-00969-0
- Cui, R., Liu, H., and Zhang, C. (2017). Recurrent convolutional neural networks for continuous sign language recognition by staged optimization. 1610–1618. doi:10.1109/CVPR.2017.175
- Damianova, N., and Berrezueta-Guzman, S. (2025). Serious games supported by virtual reality-literature review. *IEEE Access* 13, 38548–38561. doi:10.1109/access.2025.3544022
- Darejeh, A. (2023). Empowering education through erp: a customizable educational vr escape room platform. 764–766. doi:10.1109/ISMAR-Adjunct60411.2023.00166
- Economou, D., Russi, M., Doumanis, I., Mentzelopoulos, M., Bouki, V., and Ferguson, J. (2020). Using serious games for learning british sign language combining video, enhanced interactivity, and vr technology. *JUCS - J. Univers. Comput. Sci.* 26, 996–1016. doi:10.3897/jucs.2020.053
- Elhagry, A., and Elrayes, R. G. (2021). Egyptian sign language recognition using cnn and lstm. arXiv preprint arXiv:2107.13647
- Ferreira Brega, J. R., Rodello, I. A., Colombo Dias, D. R., Martins, V. F., and de Paiva Guimarães, M. (2014). "A virtual reality environment to support chat rooms for hearing impaired and to teach brazilian sign language (libras)," in 2014 IEEE/ACS 11th International Conference on Computer Systems and Applications (AICCSA), Doha, Qatar, 10–13 November 2014 (IEEE), 433–440.
- Goyal, K. (2023). Indian sign language recognition using mediapipe holistic. arXiv preprint arXiv:2304.10256
- Gunathadi, G., Atmantomi, D., Kirana Anggarani, F., and Anggrellangi, A. (2024). Virtual reality as a solution: meeting the needs of deaf students with digital sign language interpreters. *Edelweiss Appl. Sci. Technol.* 8, 1189–1199. doi:10.55214/25768484.v8i5.1823
- Gupta, S., Bagga, S., and Sharma, D. (2020). Hand gesture recognition for human computer interaction and its applications in virtual reality. 85–105. doi:10.1007/978-3-030-35252-3_5
- Hore, S., Chatterjee, S., Santhi, V., Dey, N., Ashour, A. S., Balas, V. E., et al. (2017). "Indian sign language recognition using optimized neural networks," in Information Technology and Intelligent Transportation Systems: Volume 2, Proceedings of the 2015 International Conference on Information Technology and Intelligent Transportation Systems ITITS 2015, Xi'an China, December 12–13, 2015 (Springer), 553–563.
- Horváth, I. (2020). "Personalized learning opportunity in 3d vr," in 2020 11th IEEE International Conference on Cognitive Infocommunications (CogInfoCom) (IEEE).
- Huang, J., Zhou, W., Zhang, Q., Li, H., and Li, W. (2018). Video-based sign language recognition without temporal segmentation. 32. doi:10.1609/aaai.v32i1.11903
- Humphries, T., Kushalnagar, P., Mathur, G., Napoli, D. J., Padden, C., Rathmann, C., et al. (2012). Language acquisition for deaf children: reducing the harms of zero tolerance to the use of alternative approaches. *Harm Reduct. J.* 9, 16. doi:10.1186/1477-7517-9-16
- Jiang, D., Li, G., Sun, Y., Kong, J., and Tao, B. (2019). Gesture recognition based on skeletonization algorithm and cnn with asl database. *Multimedia Tools Appl.* 78, 29953–29970. doi:10.1007/s11042-018-6748-0
- Klima, E. S., and Bellugi, U. (1979). *The signs of language*. Harvard University Press.
- Lillo-Martin, D., and Henner, J. (2021). Acquisition of sign languages. *Annu. Rev. linguistics* 7, 395–419. doi:10.1146/annurev-linguistics-043020-092357
- Liu, Y., Nand, P., Hossain, M. A., Nguyen, M., and Yan, W. (2023). Sign language recognition from digital videos using feature pyramid network with detection transformer. *Multimedia Tools Appl.* 82, 21673–21685. doi:10.1007/s11042-023-14646-0
- Luna, S. M. (2023). Dhh people in co-located collaborative multiplayer ar environments. 344–347. doi:10.1145/3573382.3616039
- Luna, S. M., Tigwell, G. W., Papangelis, K., and Xu, J. (2023). "Communication and collaboration among dhh people in a co-located collaborative multiplayer ar environment," in Proceedings of the 25th International ACM SIGACCESS Conference on Computers and Accessibility, NY, New York, USA, October 22 - 25, 2023 (IEEE), 1–5.
- Marougkas, A., Troussas, C., Krouska, A., and Sgouropoulou, C. (2023). Crafting immersive experiences: a multi-layered conceptual framework for personalized and gamified virtual reality applications in education. 230–241. doi:10.1007/978-3-031-44097-7_25
- Mazhari, A., Esfandiari, P., and Taheri, A. (2022). "Teaching iranian sign language via a virtual reality-based game," in 2022 10th RSI International Conference on Robotics and Mechatronics (ICRoM), Tehran, Iran, Islamic Republic of, 22–24 November 2022 (IEEE), 146–151.
- Monisha, H., Manish, B., and Ranjini Ravi Iyer, S. J. (2023). *Sign language detection and classification using hand tracking and deep learning in real-time*, 2395–0056.
- Morford, J. P., and Hänel-Faulhaber, B. (2011). Homesigners as late learners: connecting the dots from delayed acquisition in childhood to sign language processing in adulthood. *Lang. Linguistics Compass* 5, 525–537. doi:10.1111/j.1749-818x.2011.00296.x
- Newport, E. L. (1988). Constraints on learning and their role in language acquisition: studies of the acquisition of American sign language. *Lang. Sci.* 10, 147–172. doi:10.1016/0388-0001(88)90010-1
- Novaliendry, D., Budayawan, K., Auv, R., Fajri, B., and Huda, Y. (2023). Design of sign language learning media based on virtual reality. *Int. J. Online Biomed. Eng. (ijOE)* 19, 111–126. doi:10.3991/ijoe.v19i16.44671
- Ortiz, M., Chiluiza, K., and Valcke, M. (2017). Gamification and learning performance: a systematic review of the literature
- Parmaxi, A. (2020). Virtual reality in language learning: a systematic review and implications for research and practice. *Interact. Learn. Environ.* 31, 172–184. doi:10.1080/10494820.2020.1765392
- Qi, T. D., Cibrian, F., Raswan, M., Kay, T., Camarillo-Abad, H., and Wen, Y. (2024). Toward intuitive 3d interactions in virtual reality: a deep learning-based dual-hand gesture recognition approach. *IEEE Access* PP, 67438–67452. doi:10.1109/ACCESS.2024.3400295
- Quandt, L., Lamberton, J., Willis, A., Wang, J., Weeks, K., Kubicek, E., et al. (2020). Teaching asl signs using signing avatars and immersive learning in virtual reality. doi:10.1145/3373625.3418042

- Quinto-Pozos, D. (2011). Teaching american sign language to hearing adult learners. *Annu. Rev. Appl. Linguistics* 31, 137–158. doi:10.1017/s0267190511000195
- Ramos-cosi, S., Cardenas Pineda, C.-P., Llulluy-Nuñez, D., and Alva-Mantari, A. (2023). Development of 3d avatars for inclusive metaverse: impact on student identity and satisfaction using agile methodology, vrchat platform, and oculus quest 2. *Int. J. Eng. Trends Technol.* 71, 1–14. doi:10.14445/22315381/IJETT-V71I7P201
- Rho, E., Chan, K., Varoy, E. J., and Giacaman, N. (2020). An experiential learning approach to learning manual communication through a virtual reality environment. *IEEE Trans. Learn. Technol.* 13, 477–490. doi:10.1109/TLT.2020.2988523
- Sabbella, S. R., Kaszuba, S., Leotta, F., Serrarens, P., and Nardi, D. (2024). Evaluating gesture recognition in virtual reality. arXiv preprint arXiv:2401.04545
- Sagayam, K. M., and Hemanth, D. J. (2017). Hand posture and gesture recognition techniques for virtual reality applications: a survey. *Virtual Real.* 21, 91–107. doi:10.1007/s10055-016-0301-0
- Schick, B., Marschark, M., and Spencer, P. (2010). “The development of american sign language and manually coded english systems,” in *Oxford handbook of deaf studies*, 229–240.
- Senthamarai, S. (2018). Interactive teaching strategies. *J. Appl. Adv. Res.* 3, 36–S38. doi:10.21839/jaar.2018.v3iS1.166
- Shaw, A., Wünsche, B., Mariono, K., Ranveer, A., Xiao, M., Hajika, R., et al. (2023). Jengasl: a gamified approach to sign language learning in vr. *J. WSCG* 31, 34–42. doi:10.24132/JWSCG.2023.4
- Shi, L., Cristea, A. I., Hadzidedic, S., and Dervishalidovic, N. (2014). “Contextual gamification of social interaction – towards increasing motivation in social e-learning,” in *Advances in web-based learning – ICWL 2014*. Editors E. Popescu, R. W. H. Lau, K. Pata, H. Leung, and M. Laanpere (Cham: Springer International Publishing), 116–122.
- Shohieb, S. (2019). A gamified e-learning framework for teaching mathematics to arab deaf students: supporting an acting arabic sign language avatar. *Ubiquitous Learn. An Int. J.* 12, 55–70. doi:10.18848/1835-9795/CGP/v12i01/55-70
- Stokoe, W. C., Jr. (2005). Sign language structure: an outline of the visual communication systems of the american deaf. *J. deaf Stud. deaf Educ.* 10, 3–37. doi:10.1093/deafed/eni001
- Suh, A., Wagner, C., and Liu, L. (2016). Enhancing user engagement through gamification. *J. Comput. Inf. Syst.* 58, 204–213. doi:10.1080/08874417.2016.1229143
- Teófilo, M., Lourenço, A., Postal, J., and Lucena Jr, V. (2018). Explor virtual real enable deaf or hard hear accessibility live theaters a case study. 132–148. doi:10.1007/978-3-319-92052-8_11
- Tju, T. E. E., and Shalih, M. (2024). Hand sign interpretation through virtual reality data processing. *J. Ilmu Komput. Dan. Inf.* 17, 185–194. doi:10.21609/jiki.v17i2.1280
- Ur Rehman, M., Ahmed, F., Khan, M., Tariq, U., Alfouzan, F., Alzahrani, N., et al. (2022). Dynamic hand gesture recognition using 3d-cnn and lstm networks. *Comput. Mater. Contin.* 70, 4675–4690. doi:10.32604/cmc.2022.019586
- Wang, J., Ivrisimtzis, I., Li, Z., and Shi, L. (2024a). “Comparative efficacy of 2d and 3d virtual reality games in American sign language learning,” in 2024 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW), Orlando, FL, USA, 16–21 March 2024 (IEEE), 875–876.
- Wang, J., Ivrisimtzis, I., Li, Z., and Shi, L. (2024b). The impact of 2d and 3d gamified vr on learning american sign language. arXiv preprint arXiv:2405.08908
- Wang, J., Ivrisimtzis, I., Li, Z., and Shi, L. (2024c). “Impact of personalised ai chat assistant on mediated human-human textual conversations: exploring female-male differences,” in Companion Proceedings of the 29th International Conference on Intelligent User Interfaces, New York, NY, USA (Association for Computing Machinery), 78–83.
- Wen, H., Xu, Y., Li, L., Ru, X., Wu, Z., Fu, Y., et al. (2024). “Enhancing sign language teaching: a mixed reality approach for immersive learning and multi-dimensional feedback,” in 2024 IEEE International Conference on Systems, Man, and Cybernetics (SMC) (IEEE) (IEEE), 740–745.
- Wu, Y., Zheng, B., and Zhao, Y. (2018). “Dynamic gesture recognition based on lstm-cnn,” in 2018 chinese automation congress (CAC), 2446–2450. doi:10.1109/CAC.2018.8623035
- Zhang, W., and Wang, J. (2019). “Dynamic hand gesture recognition based on 3d convolutional neural network models,” in 2019 IEEE 16th International Conference on Networking, Sensing and Control (ICNSC), Banff, AB, Canada, 09–11 May 2019 (IEEE), 224–229.
- Zhang, Y., Shi, L., Wu, Y., Cheng, K., Cheng, J., and Lu, H. (2020). Gesture recognition based on deep deformable 3d convolutional neural networks. *Pattern Recognit.* 107, 107416. doi:10.1016/j.patcog.2020.107416
- Zhu, G., Zhang, L., Shen, P., and Song, J. (2017). Multimodal gesture recognition using 3d convolution and convolutional lstm. *IEEE Access* PP, 4517–4524. doi:10.1109/ACCESS.2017.2684186