Received July 1, 2025, August 29, 2025, date of publication September 1, 2025.

Digital Object Identifier 10.21608/ijaici.2025.422165.1016

Enhancing Educational Videos for ADHD Learners: A Review of Multimedia Design and Deep Learning Frameworks

Alshefaa Emam[1], Eman k. Elsayed[2,3], Mai K. Galab[1]

1 Faculty of Computers & Artificial Intelligence - BFCAI, Benha University, Benha, Egypt

2 Faculty of Science, Al-Azhar University (Girls)

3Department of Computer Science, College of Information Technology, Misr University for Science & Technology (MUST), 6th of October City 12566, Egypt

Corresponding author: Alshefaa Emam (e-mail: Shefaaemam97@gmail.com).

ABSTRACT In educational settings, students with Attention-Deficit/Hyperactivity Disorder often experience difficulties with maintaining focus and effectively interacting with instructional content. Video-based training has emerged as a promising approach to address these challenges, particularly as digital media continues to expand in today's classrooms. Such videos can provide children with attention-deficit/hyperactivity disorder adaptable, engaging, and motivating learning opportunities through the integration of multimedia elements and adaptive design techniques. The effectiveness of educational videos for these learners is influenced by five key factors examined in this study: video length, variety of visual objects, audio quality, visual clarity, and instructor hand movements. A review of recent research indicates that Hand Gesture Recognition technologies, multimodal cues, and high-quality audiovisual content can significantly enhance learner engagement, comprehension, and attention. Existing research, however, frequently addresses these elements separately and lacks a systematic framework that integrates them from a technical and educational perspective. This study fills this gap by putting forward a research-based, interpretable framework that combines these elements to create inclusive, successful instructional videos created especially for students with attention-deficit/hyperactivity disorder.

INDEX TERMS Deep Learning, Attention-Deficit/Hyperactivity Disorder, Educational Video Suitability, Hand Movement.

1. INTRODUCTION

With instructional videos emerging as one of the most popular digital resources, the quick expansion of online learning environments has changed the distribution and accessibility of educational content [1], [2]. Although these videos offer flexibility, multimodal interaction, and a worldwide audience, not all students benefit equally from them. Students with Attention-Deficit/Hyperactivity Disorder (ADHD), a neurodevelopmental disorder that affects roughly 6.8% of adults and 7.6% of children globally [3], in particular, frequently have significant difficulties when it comes to participating in and reaping the benefits of video-based training. Their capacity to maintain focus, absorb information, and remember knowledge can be weakened by symptoms such impulsivity, hyperactivity, and inattention [4], [5]. According to research, learners with ADHD may be further hampered by instructional videos that lack adequate diversity, interactive cues, or suitable pacing [6], [7], [8]. Additionally, excessive or badly structured visual and audio stimulation might reduce understanding and increase distraction [9]. Video design must therefore be carefully modified to suit neurodiverse learners. Numerous studies have examined many aspects over the last ten years, including multimedia components, video duration, instructor hand gestures, and audio-visual quality [10], [11]. More recently, educational videos have been evaluated and categorised using sophisticated computational techniques like explainable artificial intelligence (XAI) and deep learning that take into account the cognitive needs of students with ADHD [12]. There are still large gaps in spite of these advancements. A large portion of the material currently in publication is based on manual, small-scale assessments that are not generally applicable, while more comprehensive research on online learning frequently ignores difficulties unique to ADHD [13]. Furthermore, there is very limited use of deep learning models for assessing educational content for ADHD, despite their impressive performance in related fields like image recognition and video analysis [14]. This paper makes two major contributions to alleviate these limitations. It begins by reviewing current studies on video-based learning for students with ADHD, with an emphasis on five factors: the duration of the video, the variety of objects, the audio quality, the visual clarity, and the hand gestures of the instructor. Secondly, it explores how deep learning techniques can be used to automatically recognise video features that enhance learning outcomes and engagement for neurodiverse students. In order to give a thorough review of the topic and suggest future research possibilities, this book attempts to bridge insights from educational psychology, instructional design, and artificial intelligence. This paper's remaining sections are organised as follows. Section 2 examines the related work on video-based instruction for ADHD students, highlighting important results and major research needs. Section 3 provides a comprehensive discussion of the evaluated literature on HGR methods and the applicability of educational videos for ADHD, followed up by a summary table. The paper's conclusion and possible future study directions are provided in Section 4.

2. Related Work

Due to the increasing prevalence of ADHD worldwide and the expanding significance of digital media in education, there has been a growing interest in using instructional videos to help students with ADHD. Inattention, hyperactivity, and impulsivity are the hallmarks of ADHD, which affects roughly 7.6% of children and 6.8% of adults globally [3], [15], [16] . These traits frequently impair motivation and academic performance. When carefully created, instructional films provide structured, multimodal content that is suited to a variety of attention patterns in order to overcome these issues. The incorporation of interactive components that maintain focus is a crucial component. A subfield of human-computer interaction called hand gesture recognition (HGR) serves as one illustration. In recent years, developments in deep learning and computer vision have improved engagement through HGR, which offers visual anchors that direct attention, support spoken content, and explain challenging concepts all of which are especially helpful for students who struggle with attention management.

Audio-visual quality is equally important. Clear images and well-balanced audio, according to research, assist ADHD children avoid distractions, nevertheless low-quality sound, background noise, or inadequate illumination cause overload and disengagement. Audio-visual content that has been optimized promotes comprehension, focus, and retention.

The variety of visual objects is another crucial factor. Learning is enhanced and visual comprehension is supported when on-screen text, diagrams, animations, and other components are used. Balanced design is crucial since too many or badly organized images can overwhelm ADHD students. In order to assess their effects, this paper uses machine learning and deep learning to examine five fundamental multimedia dimensions: teacher hand movements, video length, object diversity, audio quality, and visual quality. This paper supports inclusive digital education by filling up theoretical and empirical gaps and directing the creation of adaptive learning platforms and video content that is suitable for people with ADHD.

2.1 Video-Based Classification for ADHD Students

Significant academic obstacles, such as subpar grades, decreased standardized test scores, and higher rates of grade retention, are faced by students with ADHD. According to research, in order to lessen these difficulties, specialized teaching methods are required. Ten primary school students with ADHD and their instructors participated in semi-structured interviews with McDougal et al., who identified classroom tactics included dividing work into digestible portions and using organizing aids like homework folders. However these approaches are frequently generic rather than ADHD-specific, suggesting the need for more focused therapies [17]. In a similar vein, Lovett & Nelson carried out a systematic review of accommodations for kids and teens with ADHD, including extended test duration, read-aloud choices, small-group testing, color contrast tools, and calculators. They discovered that although these accommodations are frequently utilized, they frequently lack empirical support specific to ADHD and exhibit inconsistent efficacy [18]. Although they don't particularly address video-based therapies, these studies highlight the value of individualized teaching techniques. Video-based education presents a viable way to engage students with ADHD. Students with ADHD benefit from multimodal, interactive, and succinct content to help them with their attention issues, and videobased learning is a promising way to engage them. During the COVID-19 pandemic, Levenberg & Abu Reesh investigated how 12 university students with ADHD perceived learning from recorded lectures. They discovered that video recordings give students a sense of control over time, place, and pace, allowing them to pause, rewind, or change playback speed to suit their needs [19]. Although there are no systematic frameworks to categorise educational videos based on ADHD-specific criteria like attention retention or emotional regulation, these studies indicate that educational videos can be created with features like visual cues, brief segments, and interactive elements to decrease cognitive load and increase engagement for ADHD students. The PEMAT technique for audiovisual content in Japanese patient education, for instance, was translated and verified by Furukawa et al., demonstrating its ongoing applicability in assessing the actionability and understandability of multimedia materials [22]. The PEMAT technique for audiovisual content in Japanese patient education, for instance, was translated and verified by Furukawa et al., demonstrating its ongoing applicability in assessing the actionability and understandability of multimedia materials [20], To counteract disinformation strategies, Ji et al. presented self-reflection techniques in big language models, which could be modified to guarantee the accuracy of classified instructional videos [21], Thus, the literature reveals a number of challenges: The variety of ADHD symptoms (e.g., inattentive vs. hyperactive-impulsive subtypes) necessitates adaptive classification systems based on individual differences, a gap not addressed in current literature [23]. However, rather than being comparative, a large portion of the examined literature is descriptive. Previous research frequently presents tools or methodologies separately without conducting a thorough evaluation. Seldom are important elements like datasets, performance metrics, and study constraints compared. For instance, while some study emphasises aspects like interactive elements or visual signals, there isn't much data on which tactics are most effective for students with ADHD. It is challenging to generalise results or create precise design guidelines for video-based training due to the absence of organised comparison.

2.2 Hand Gesture Recognition

HGR is a contactless and user-friendly way of interaction that uses cameras to record and interpret motions. In their thorough analysis of 90 studies, Ameur et al. highlighted the Leap Motion Controller (LMC) for accurate gesture tracking in robotics and medical applications, pointing out both its advantages and disadvantages in terms of data collection and calibration [24]. Similar to this Khan et al. divided vision-based methods into four categories: deep learning, motion analysis, skeleton-based modelling, and skin colour segmentation. They also identified problems such as background noise and illumination variation [25], which are especially important in real-world situations where reliable segmentation is crucial. The goal of recent developments is to enhance real-time performance. For example, Garcia et al. proposed RGB-S, which avoids computationally costly techniques like optical flow and achieves high accuracy on the IPN Hand dataset by combining RGB frames with hand segmentation masks via Temporal Segment Networks (TSN) and ResNet-50 [26]. In order to enable biomimetic learning for humanoid robots, Olikkal et al. used the MediaPipe architecture to recognise 33 static American Sign Language (ASL) movements in real-time using a single RGB camera. [27], despite the fact that their focus on static movements restricts their use in dynamic situations. Linardakis et al. examining visual-input-based HGR covering RGB, depth, and video data and summarised current trends and challenges like occlusion, generalisation across users, and computational efficiency for real-time deployment, highlight the need for current comprehensive surveys. For accurate detection, sensor-based HGR techniques that record hand movements or muscle signals include datagloves and surface electromyography (sEMG). Standardising feature extraction (e.g., RMS, MAV) and classification (e.g., SVM, ANN) procedures in their review of 65 sEMG-based research, pointing out issues with overfitting and reproducibility. [28]. Although signal noise is still an issue, Zhang et al. described sophisticated applications like fingertip force estimation and catalogued datasets like Ninapro and CapgMyo. [29]. Another sensor-based method is Datagloves: In order to recognise 24 static and 16 dynamic ASL motions, Faisal et al. developing a low-cost dataglove with flex sensors and inertial measurement units and presented a Spatial Projection Image approach. Despite having a small dataset, their parallel-path neural network performed better than conventional ML and CNN techniques [30]. In their comparison of vision-based and sensor-based approaches, Kim et al. emphasised the computational intricacy of datagloves and the significance of reliable algorithms such as deep neural networks and Hidden Markov Models [31]. Applications of HGR include robots, virtual reality (VR), sign language translation, and human–computer interface. Using color-space segmentation and machine learning, Smith et al. created a two-phase British Sign Language (BSL) recognition system that is restricted to five BSL alphabets and is aimed at the speech-impaired community[32]. Chen et al. compared systems based on gloves, cameras, radar, and EMG in VR situations and recommended including tactile feedback for immersive experiences [33]. Although scalability in complex situations is still a challenge, it investigated biomimetic learning for human–robot collaboration in the field of robotics[27]. whereas vision-based techniques are hindered by occlusion, background noise, and variations in light [25], Important research gaps still exist: sEMG-based techniques suffer from signal noise and variability [28], [34]. The amount, diversity, and recentness of datasets across methodologies are frequently lacking [26], [35]. Lightweight models are required for real-time implementation in devices with limited resources, and generalisability is limited by concentrating on small gesture sets such as ASL or BSL [30], [32]. Lastly, comparing different methodologies is made more difficult by the lack of standardised datasets and evaluation measures [28], [31]. Overall, earlier studies on HGR show notable advancements, However they are still dispersed. Sensor-based approaches have problems with noise and repeatability, whereas vision-based approaches have problems with occlusion and lighting. Generalisation is challenging because these works are rarely compared using uniform datasets or evaluation metrics. Furthermore, generalisability is diminished by the majority of studies' emphasis on small gesture sets. The necessity for a methodical framework that unifies research results and directs future HGR implementations in practical settings is highlighted by this gap.

2.3 Video and Audio Quality

For children with ADHD, who are extremely sensitive to sensory distractions and need clear, captivating stimuli, the quality of the audio and video is crucial in deciding how successful the educational material is. A. Yeung, decleared background noise and other poor audio quality can seriously hinder speech processing and comprehension [36], For children with ADHD, who are extremely sensitive to sensory distractions and need clear, captivating stimuli, the quality of the audio and video is crucial in deciding how successful the educational material is. Background noise and other poor audio quality can greatly hinder speech processing and understanding, highlighting the need of clean audio in reducing distractions and promoting auditory processing. Lui et al., using the Patient Education Materials Assessment Tool (PEMAT-A/V), an analysis of 100 TikTok videos about ADHD showed that high-quality audio-visual content created by medical professionals was more accurate and captivating, highlighting the importance of quality in thwarting false information and maintaining focus. [37]. L. Jiang, has been demonstrated that short, visually appealing psychoeducational videos with crystal-clear audio increase accessibility and engagement for young people with ADHD from a variety of language backgrounds. [38]. There is no study on optimising audio-visual quality measures, like resolution, audio clarity, and captioning, for ADHD-specific cognitive and sensory demands, despite these discoveries [20], [36], he majority of research focusses on user-generated content or generic accommodations instead of organized instructional videos that are adapted to the learning objectives of ADHD [37], [38], and existing designs do not sufficiently address the variability of ADHD symptoms, which call for flexible content forms [23]. All things considered, although earlier research emphasises the significance of audio and video clarity for sustaining focus, these findings are rarely translated into systematic design principles or assessment frameworks. This disparity emphasises the necessity of using standardised criteria to evaluate the audiovisual quality of instructional videos, especially those geared towards students with ADHD, in order to make sure that multimedia content reduces distractions and enhances understanding.

2.4 Object Detection

Since ADHD children can be both hyper focused and distracted, a variety of objects in movies, such as props, graphics, or on-screen elements, are essential for keeping their attention. H. Zhu, R. well-chosen visual components and regulated variability promote focus, however excessive or unnecessary visuals increase distraction and cognitive overload, according to a recent study examining video accessibility tactics for ADHD viewers [39]. Similarly, A. Shrivastava, was shown that decreasing clutter, emphasising important visual elements, and simplifying backdrops greatly enhance viewability and lessen distraction in Focus View, a customisable video interface made for people with ADHD [40]. Validated through a nine-month experimental study with children with ADHD, design principles based on neurodiversity-informed visual design demonstrated that incorporating visually distinct elements, such as clear focus points, simplified icons, and limited clutter, gradually improves attention and engagement. Despite these advancements, there are still important gaps in the literature: few empirical studies have been conducted to quantify the precise effects of visual object quantity, type, or complexity on attention and learning in ADHD learners [39], [40], and the majority of studies concentrate on general accessibility principles rather than educational videos specifically designed for ADHD [40], The range of ADHD symptoms highlights the need for individualized visual design approaches, which are still little understood [23], and there aren't many systematic frameworks for creating movies with a variety of visual objects that are targeted at ADHD [40]. In order to maximize the visual components of instructional films for students with ADHD, these gaps underscore the pressing need for evidence-based design recommendations. Ultimately, previous research has recognised the influence of visual items on attention, however it has not provided systematic frameworks for design or assessment that are specific to students with ADHD. This reveals a crucial gap: in order to develop evidence-based, ADHD-specific guidelines for incorporating visual elements into instructional videos, future research must go beyond generic accessibility principles. Attempts like these would guarantee that object variety promotes concentration without causing distraction or cognitive overload.

3. Discussion

The reviewed literature's recurrent themes are shown in Table 1. A technological viewpoint emphasises HGR and computer vision methods. On the other hand, an educational viewpoint stresses the appropriateness of videos, taking into account elements like length, visual variety, and audio/visual quality. The current research can be roughly categorised using these two viewpoints. Despite the abundance of tools and technological developments in computer vision, there are still few systematic frameworks available for classifying instructional videos. More significantly, only a small percentage of these frameworks are made especially to meet the needs of students with ADHD. This disparity emphasises the need for strategies that incorporate both technical and instructional factors.The table also demonstrates how accuracy is the primary focus of reported performance indicators, which frequently rely on tiny or unbalanced datasets and make little use of standardised evaluation criteria. This limits the findings' capacity to be compared and generalised. Furthermore, rather than including the visual, auditory, and temporal modalities into a holistic multimodal model, the majority of research handle them separately. The lack of explainable outputs, a crucial element for the effective use of such systems in educational settings, is another significant drawback. Practically speaking, our results highlight the necessity of multimodal categorisation frameworks that integrate temporal, visual, and aural inputs, backed by lightweight models that may be implemented in actual learning settings. In order to give educators interpretable explanations for system judgements, explainable AI (XAI) must be incorporated. In this way, the table acts as a road map: the most promising avenue for further study is to incorporate the powerful technical capabilities of computer vision into an interpretable and educationally based framework specifically designed for students with ADHD. The review identifies a number of gaps in the existing literature that need to be filled:

- Absence of integrated frameworks: Previous research addresses educational elements (like video appropriateness for ADHD) and technical aspects (like HGR and computer vision) alone, without any kind of systematic integration.

- Over-reliance on limited measurements Accuracy is the primary focus of reported performance metrics, which frequently use unstandardised evaluation procedures and small or unbalanced datasets.

- Lack of multimodal models: Most research examines visual, auditory, or temporal modalities separately rather than in a whole multimodal framework designed for students with ADHD.

- Limited interpretability: Only a small number of methods use explainable AI, which limits the results' transparency and prevents their use in actual learning settings.

These combined limitations highlight the need for more research that combines pedagogical design and technological innovation to make sure that classification systems are useful and instructive for students with ADHD.

TABLE 1. Summary of Reviewed Studies on ADHD Educational Video Suitability and HGR Techniques

Research Number	Dataset Used	Methodology	Results	Relevance to ADHD	Identified Gaps
Polanczyk et al., 2014 **[4]**	10 primary students with ADHD + teachers	Semi-structured interviews	Identified strategies (task chunking, homework folders)	Highlights classroom accommodations	General strategies; not ADHD-specific interventions
Becker et al., 2020 **[5]**	Multiple studies	Systematic review	Accommodations (extended time, small groups, etc.) show mixed effectiveness	Supports flexible learning approaches	Lack of ADHD-specific empirical support
Mayer, Fiorella & Stull, 2020 **[6]**	12 university students with ADHD	Interviews and surveys	Recorded lectures provide flexibility in pace, location, timing	Video recordings support ADHD self-paced learning	No ADHD-specific video classification framework
Dargue, Sweller & Jones, 2019 **[7]**	Japanese patient education videos	Translation + validation study	PEMAT tool validated for audiovisual content	Shows tools for evaluating multimedia clarity	Not tailored to ADHD educational video needs
Li et al., 2022 **[8]**	Large language models	AI self-reflection experiment	Reduced misinformation in AI responses	Adaptable to ensure accurate video classification	Not applied to ADHD or video-based learning
Holmes, Bialik & Fadel, 2019 **[9]**	Theoretical study	Literature insights	ADHD symptoms vary (inattentive vs hyperactive)	Highlights need for adaptive frameworks	Current systems don't address ADHD subtypes
Oudah, Al-Naji & Chahl, 2020 **[24]**	Leap Motion datasets	Review of Leap Motion in HGR	Achieved 85–95% accuracy - Effective in robotics and	Emphasizes Leap Motion for precise vision-based tracking	Occlusion and calibration issues; limited generalizability

			medical applications		
Benitez-Garcia et al., 2021 **[25]**	NUS Hand Posture, ASL Fingerspelling	Systematic review of vision-based techniques	CNNs and depth sensors achieved 70–95% accuracy - Challenges: illumination, background noise	Categorizes vision-based methods (e.g., skin segmentation, deep learning); highlights real-world challenges	Limited dataset diversity; lack of standardized evaluation metrics
Olikkal et al., 2024 **[26]**	IPN Hand	Semantic segmentation using TSN, ResNet-50	Accuracy improved by 10% over baseline CNNs - Real-time processing via RGB-S	Advances real-time HGR with RGB and segmentation; avoids costly optical flow	Small dataset size; limited to static gestures
Jaramillo-Yánez, Benalcázar & Mena-Maldonado, 2020 **[27]**	Custom ASL dataset	Biomimetic learning using MediaPipe	88% accuracy for 33 static ASL gestures - Supports human–robot collaboration	Real-time vision-based HGR for robotics using a single RGB camera	Limited to static gestures; small dataset size
Wang, 2024 **[32]**	Custom BSL dataset	Color-space segmentation + ML	90% accuracy for 5 British Sign Language alphabets - Reduced performance in dynamic gestures	Vision-based BSL recognition for speech-impaired users	Limited to a few gestures; poor generalizability
Ni et al., 2024 **[33]**	VR-based datasets	Review of AI in VR-based HGR	Achieved 90–95% accuracy - Hardware choice impacts performance	Compares vision- and sensor-based HGR in VR; advocates tactile feedback	Resource-intensive; lack of standardized metrics
Lemel, Shalev, Nitsan & Ben-David, 2023 **[35]**	IPN Hand, DHG-14/28	Survey of vision-based HGR (RGB, depth, video)	CNNs and LSTMs achieved 90–98% accuracy - Dataset diversity found to be critical	Comprehensive review of visual-input HGR; addresses occlusion and generalization	Lack of standardized datasets; real-time deployment challenges
Yeung, Ng & Abi-Jaoude, 2022 **[36]**	24 young adults with ADHD	Experimental study (eye-tracking, neuroimaging)	Poor audio quality impairs speech processing - Background noise reduces ADHD comprehension	Emphasizes clear audio to minimize sensory distractions for ADHD	Limited to audio; no comprehensive video quality metrics
Lui et al., 2025 **[37]**	100 TikTok videos on ADHD	Cross-sectional content analysis (PEMAT-A/V)	High-quality AV content more accurate/engaging - Healthcare professional videos outperformed others	Highlight's role of high-quality audio-visuals in engaging ADHD students	Focus on user-generated content; no ADHD-specific quality framework

Jiang et al., 2025 **[38]**	Youth with ADHD	Development/evaluation of online resources	Short, clear AV psychoeducational videos improved engagement	Short, high-quality videos enhance accessibility for ADHD youths	Limited to general psychoeducation; no systematic quality classification

4.Conclusion

This review emphasised two complementary research streams: (i) educational research that focusses on the pedagogical and cognitive aspects of designing videos that are suitable for people with ADHD, and (ii) technological advancements in computer vision and HGR that offer powerful tools for modelling behavioural and visual features. The development of adaptable, captivating, and customised video-based learning for children with ADHD is made possible by bridging these areas. Future studies should focus on developing multimodal and balanced datasets specifically for ADHD learners, integrating explainable AI (XAI) for clear and understandable decision-making, and developing conceptual frameworks that methodically integrate educational suitability factors with gesture-based and visual recognition techniques in order to further this field. Additionally, to evaluate how well these strategies improve accessibility, interaction, and learning outcomes for neurodiverse children, empirical validation in actual classroom settings is required.

REFERENCE

[1] N. Al Mudawi et al., "Innovative healthcare solutions: robust hand gesture recognition of daily life routines using 1D CNN," Front Bioeng Biotechnol, vol. 12, p. 1401803, 2024, doi: 10.3389/fbioe.2024.1401803 .

[2] H. Crompton and D. Burke, "Artificial intelligence in higher education: the state of the field," International journal of educational technology in higher education, vol. 20, no. 1, p. 22, 2023, doi: 10.1186/s41239-023-00392-8 .

[3] S. V Faraone et al., "The world federation of ADHD international consensus statement: 208 evidence-based conclusions about the disorder," Neurosci Biobehav Rev, vol. 128, pp. 789–818, 2021, doi: 10.1016/j.jaac.2023.07.191 .

[4] G. V Polanczyk, E. G. Willcutt, G. A. Salum, C. Kieling, and L. A. Rohde, "ADHD prevalence estimates across three decades: an updated systematic review and meta-regression analysis," Int J Epidemiol, vol. 43, no. 2, pp. 434–442, 2014, doi: 10.1093/ije/dyt261 .

[5] S. P. Becker et al., "Remote learning during COVID-19: Examining school practices, service continuation, and difficulties for adolescents with and without attention-deficit/hyperactivity disorder," Journal of Adolescent Health, vol. 67, no. 6, pp. 769–777, 2020, doi: 10.1016/j.jadohealth.2020.09.002 .

[6] R. E. Mayer, L. Fiorella, and A. Stull, "Five ways to increase the effectiveness of instructional video," Educational Technology Research and Development, vol. 68, no. 3, pp. 837–852, 2020, doi: 10.1007/s11423-020-09749-6 .

[7] N. Dargue, N. Sweller, and M. P. Jones, "When our hands help us understand: A meta-analysis into the effects of gesture on comprehension.," Psychol Bull, vol. 145, no. 8, p. 765, 2019, doi: 10.1037/bul0000202 .

[8] X. Li, Y. Dong, Y. Jiang, and G. A. Ogunmola, "Analysis of the teaching quality of college ideological and political education based on deep learning," Journal of Interconnection Networks, vol. 22, no. Supp02, p. 2143002, 2022, doi: 10.1142/s0219265921430027 .

[9] W. Holmes, M. Bialik, and C. Fadel, Artificial intelligence in education promises and implications for teaching and learning. Center for Curriculum Redesign, 2019.

[10] G. J. DuPaul, S. W. Evans, J. A. Mautone, J. S. Owens, and T. J. Power, "Future directions for psychosocial interventions for children and adolescents with ADHD," Journal of Clinical Child & Adolescent Psychology, vol. 49, no. 1, pp. 134–145, 2020, doi: 10.1080/15374416.2019.1689825 .

[11] Z. Pi, Y. Zhang, F. Zhu, K. Xu, J. Yang, and W. Hu, "Instructors' pointing gestures improve learning regardless of their use of directed gaze in video lectures," Comput Educ, vol. 128, pp. 345–352, 2019, doi: 10.1016/j.compedu.2018.10.006 .

[12] O. Chen, J. C. Castro-Alonso, F. Paas, and J. Sweller, "Extending cognitive load theory to incorporate working memory resource depletion: Evidence from the spacing effect," Educ Psychol Rev, vol. 30, no. 2, pp. 483–501, 2018, doi: 10.1007/s10648-017-9426-2 .

[13] N. Alquwez et al., "Self-compassion influences the caring behaviour and compassion competence among saudi nursing students: A multi-university study," Nurs Open, vol. 8, no. 5, pp. 2732–2742, 2021, doi: 10.1002/nop2.848 .

[14] C. Rudin, "Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead," Nat Mach Intell, vol. 1, no. 5, pp. 206–215, 2019, doi: 10.1038/s42256-019-0048-x .

[15] S. Popit, K. Serod, I. Locatelli, and M. Stuhec, "Prevalence of attention-deficit hyperactivity disorder (ADHD): systematic review and meta-analysis," European Psychiatry, vol. 67, no. 1, p. e68, 2024.

[16] P. Song, M. Zha, Q. Yang, Y. Zhang, X. Li, and I. Rudan, "The prevalence of adult attention-deficit hyperactivity disorder: A global systematic review and meta-analysis," J Glob Health, vol. 11, p. 04009, 2021.

[17] E. McDougal, C. Tai, T. M. Stewart, J. N. Booth, and S. M. Rhodes, "Understanding and supporting attention deficit hyperactivity disorder (ADHD) in the primary school classroom: Perspectives of children with ADHD and their teachers," J Autism Dev Disord, vol. 53, no. 9, pp. 3406–3421, 2023, doi: 10.1007/s10803-022-05639-3 .

[18] B. J. Lovett and J. M. Nelson, "Systematic review: Educational accommodations for children and adolescents with attention-deficit/hyperactivity disorder," J Am Acad Child Adolesc Psychiatry, vol. 60, no. 4, pp. 448–457, 2021, doi: 10.1016/j.jaac.2020.07.891 .

[19] A. Levenberg and S. A. Reesh, "Learning from recorded lectures: Perceptions of students with ADHD," J Atten Disord, vol. 27, no. 9, pp. 960–972, 2023, doi: 10.1177/10870547231171725 .

[20] E. Furukawa et al., "Translation, cross-cultural adaptation, and validation of the Japanese version of the patient education materials assessment tool (PEMAT)," Int J Environ Res Public Health, vol. 19, no. 23, p. 15763, 2022, doi: 10.21203/rs.3.rs-909531/v1 .

[21] Z. Ji, T. Yu, Y. Xu, N. Lee, E. Ishii, and P. Fung, "Towards mitigating LLM hallucination via self reflection," in Findings of the Association for Computational Linguistics: EMNLP 2023, 2023, pp. 1827–1843, doi: 10.18653/v1/2023.findings-emnlp.123 .

[22] N. Salari et al., "The global prevalence of ADHD in children and adolescents: a systematic review and meta-analysis," Ital J Pediatr, vol. 49, no. 1, p. 48, 2023, doi: 10.1186/s13052-023-01456-1 .

[23] N. M. Bhiri, S. Ameur, I. Alouani, M. A. Mahjoub, and A. Ben Khalifa, "Hand gesture recognition with focus on leap motion: An overview, real world challenges and future directions," Expert Syst Appl, vol. 226, p. 120125, 2023, doi: 10.1016/j.eswa.2023.120125 .

[24] M. Oudah, A. Al-Naji, and J. Chahl, "Hand gesture recognition based on computer vision: a review of techniques," J Imaging, vol. 6, no. 8, p. 73, 2020, doi: 10.3390/jimaging6080073 .

[25] G. Benitez-Garcia et al., "Improving real-time hand gesture recognition with semantic segmentation," Sensors, vol. 21, no. 2, p. 356, 2021, doi: 10.3390/s21020356 .

[26] P. Olikkal, D. Pei, B. K. Karri, A. Satyanarayana, N. M. Kakoty, and R. Vinjamuri, "Biomimetic learning of hand gestures in a humanoid robot," Front Hum Neurosci, vol. 18, p. 1391531, 2024, doi: 10.3389/fnhum.2024.1391531 .

[27] A. Jaramillo-Yánez, M. E. Benalcázar, and E. Mena-Maldonado, "Real-time hand gesture recognition using surface electromyography and machine learning: A systematic literature review," Sensors, vol. 20, no. 9, p. 2467, 2020, doi: 10.3390/s20092467 .

[28] C. Wang et al., "Survey on factuality in large language models: Knowledge, retrieval and domain-specificity," arXiv preprint arXiv:2310.07521, 2023.

[29] M. A. A. Faisal, F. F. Abir, M. U. Ahmed, and M. A. R. Ahad, "Exploiting domain transformation and deep learning for hand gesture recognition using a low-cost dataglove," Sci Rep, vol. 12, no. 1, p. 21446, 2022, doi: 10.1038/s41598-022-25108-2 .

[30] M. M. Rahman, A. Uzzaman, F. Khatun, M. Aktaruzzaman, and N. Siddique, "A comparative study of advanced technologies and methods in hand gesture analysis and recognition systems," Expert Syst Appl, vol. 266, p. 125929, 2025, doi: 10.1016/j.eswa.2024.125929 .

[31] V. Chang, R. O. Eniola, L. Golightly, and Q. A. Xu, "An exploration into human–computer interaction: Hand gesture recognition management in a challenging environment," SN Comput Sci, vol. 4, no. 5, p. 441, 2023, doi: 10.1007/s42979-023-01751-y .

[32] J. Wang, "Application of artificial intelligence in hand gesture recognition with virtual reality: Survey and analysis of hand gesture hardware selection," arXiv preprint arXiv:2405.16264, 2024.

[33] S. Ni, M. A. A. Al-qaness, A. Hawbani, D. Al-Alimi, M. Abd Elaziz, and A. A. Ewees, "A survey on hand gesture recognition based on surface electromyography: Fundamentals, methods, applications, challenges and future trends," Appl Soft Comput, vol. 166, p. 112235, 2024, doi: 10.1016/j.asoc.2024.112235 .

[34] M. Linardakis, I. Varlamis, and G. T. Papadopoulos, "Survey on hand gesture recognition from visual input," arXiv preprint arXiv:2501.11992, 2025, doi: 10.1109/access.2025.3593428 .

[35] R. Lemel, L. Shalev, G. Nitsan, and B. M. Ben-David, "Listen up! ADHD slows spoken-word processing in adverse listening conditions: Evidence from eye movements," *Res Dev Disabil*, vol. 133, p. 104401, 2023.

[36] A. Yeung, E. Ng, and E. Abi-Jaoude, "TikTok and attention-deficit/hyperactivity disorder: a cross-sectional study of social media content quality," *The Canadian Journal of Psychiatry*, vol. 67, no. 12, pp. 899–906, 2022, doi: 10.1177/07067437221082854.

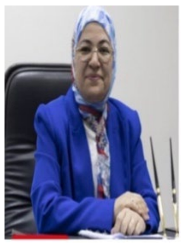
[37] J. H. L. Lui *et al.*, "Development and Dissemination of Online Evidence-Based Information for Youths with ADHD and Their Adult Supports," Evid Based Pract Child Adolesc Ment Health, pp. 1–17, 2025, doi: 10.1080/23794925.2025.2457632 .

[38] L. Jiang, W. Ko, S. Yuan, T. Shende, and S. Azenkot, "Shifting the Focus: Exploring Video Accessibility Strategies and Challenges for People with ADHD," in Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems, 2025, pp. 1–16, doi: 10.1145/3706598.3713637 .

[39] H. Zhu, R. Chen, and Y. Zhao, "FocusView: Understanding and Customizing Informational Video Watching Experiences for Viewers with ADHD," arXiv preprint arXiv:2507.13309, 2025.

[40] S. S. Phalke and A. Shrivastava, "Evaluating visual design guidelines for digital content representation: A nine-month study on enhancing attention in children with ADHD," Authorea Preprints, 2025, doi: 10.31124/advance.174314142.24494763/v1 .

Alshefaa Mohamed Emam received the B.Sc. degree in Information Technology from Misr University for Science & Technology (MUST). She is currently pursuing an M.Sc. degree in Computers and Artificial Intelligence at Benha University. She is a teaching assistant at Zewail City, School of Computational Sciences & Artificial Intelligence. Her research interests include deep learning, computer vision, and AI-driven educational systems. She has published research in the field of AI and its applications in educational content analysis. She can be contacted email at: shefaaemam97@gmail.com.



Prof. Eman K. Elsayed is the Dean of the College of Information Technology at Misr University for Science & Technology (MUST). She previously served as the Dean of the School of Computer Science at the Canadian International College. She earned her Bachelor of Science degree in Computer Science from Cairo University in 1994, followed by a Master's degree in Computer Science from Cairo University in 1999. She obtained her Ph.D. in Computer Science from Al-Azhar University in 2005 and was promoted to Professor of Computer Science in 2019. With a strong background in artificial intelligence, Prof. Elsayed has published 65 research papers across various AI domains as of March 2024. She can be contacted via email at: eman.karam@must.edu.eg.

Dr. Mai Kamal El-Din Galab received her BSc in Computer Science from the Faculty of Science, Benha University, in 2007. She completed her Master's degree in Computer Science at the Faculty of Computers & Artificial Intelligence, Benha University, in 2015. Dr. Galab earned her Ph.D. in 2022 from the same faculty. Since 2022, she has been a Lecturer in Computer Science. Her research interests are computer vision, Image Processing (Human Behavior Analysis - Video Surveillance Systems), Security (Steganography –Encryption), and Human-Computer Interaction (HCI). Dr. Galab can be contacted via email at: may.gallab@fci.bu.edu.eg.