# Review of Studies on Recognition Technologies and Their Applications Used to Assist Learning and Instruction

**Rustam Shadiev[1], Zi Heng Zhang[1], Ting-Ting Wu[2] and Yueh Min Huang[3*]**
[1]School of Education Science, Nanjing Normal University, China // [2]National Yunlin University of Science and Technology, Taiwan // [3]Department of Engineering Science, National Cheng Kung University, Taiwan //
rustamsh@gmail.com // 1040303326@qq.com // ttwu@yuntech.edu.tw // huangncku@gmail.com
[*]Corresponding author

(Submitted February 17, 2020; Revised April 6, 2020; Accepted September 9, 2020)

**ABSTRACT:** We reviewed studies on recognition technologies published in the last ten years. This review study was aimed toward identifying, appraising, selecting, and synthesizing all high quality research evidence published in the literature related to recognition technologies and on determining how they can assist learning and instruction. This study particularly focuses on summarizing the current state of knowledge in the following dimensions: (1) recognition technology and processes, (2) applications, (3) schemes, and (4) advantages and disadvantages. We reviewed seventy-two papers and identified eighteen recognition technologies. Our results showed that all of the recognition technologies under consideration featured different recognition processes and applications. In most studies, the participants were university students. Questionnaires and tests were the most frequently used data collection methods. Most studies used a group comparison as their research design. Finally, several advantages and disadvantages of the recognition technologies were identified and summarized in the papers. The most frequently cited disadvantage was a low recognition accuracy rate. Based on our results, several suggestions and implications are made for the teaching and research community.

**Keywords:** Review, Recognition technology, Learning and instruction

## 1. Introduction

Signal processing has recently become an increasingly advancing field due to rapid development of information and communication technologies. Signal processing is a branch of electrical engineering defined as a process in which information is extracted from a signal (Shih, 2010). According to van De Ville, Demesmaeker, and Preti (2017), signal processing focuses on analyzing, modifying, and synthesizing signals such as sounds, images, and biological measurements. Signal processing manipulates information content in signals and is intended to facilitate automatic recognition.

Recognition technologies are vital signal processing applications (Dong, Wang, & Kuang, 2014; Shadiev, Hwang, Chen & Huang, 2014). These technologies extract information from signals, analyze it, and then translate it into recognizable output. One example of such a technology is speech-to-text recognition. This technology receives voice input from a speaker, processes it, and then generates textual output (Shadiev et al., 2014). It has been suggested that recognition technologies can be useful with assisting learning and instruction. In the field of education, there are different types of input signals, e.g., speech, facial images, gesture images/videos, or physiological signals. The purpose of these signals is to collect data from learners and translate the data into recognizable learning data either in cognitive or affective domains. There are currently many recognition technologies, e.g., emotion recognition (Lin, Su, Chao, Hsieh, & Tsai, 2016), hand movement recognition (Wolski & Jagodziński, 2019), physical activity recognition (Lindberg, Seo, & Laine, 2016), gesture recognition (Yang & Liao, 2014), handwriting recognition (Yanikoglu, Gogus, & Inal, 2017), facial recognition (Yueh, Lin, Liu, Shoji, & Minoh, 2014), haptic recognition (Magana & Balachandran, 2017), touch recognition (Mercier, Vourloumi, & Higgins, 2017), image recognition (Hwang, Chang, Chen, & Chen, 2018), sign recognition (Ditcharoen, Naruedomkul, & Cercone, 2010), attention recognition (Chen, Wang, & Yu, 2017), biometric information recognition (Kaklauskas et al., 2010), math expression recognition (Pacheco-Venegas, López, & Andrade-Aréchiga, 2015), body position and movement recognition (Wang, Hwang, Li, Chen, & Manabe, 2019), etc. Scholars have applied such technologies to aid both learning and instruction.

There are currently many review studies on different recognition technologies. Calvo and Nummenmaa (2016) reviewed studies on facial expression recognition to explore the role of visual and emotional factors in the recognition process. The results showed that these factors contribute differently to the process; however, facial expression recognition depends on perceptual factors more than affective factors. Ghosh, Dube, and Shivaprasad (2010) synthesized results from script recognition studies focusing on different script recognition process

methodologies. Their results revealed two distinct script recognition techniques: structure-based and visual appearance-based, and two different script features for script identification: character-wise and word-wise. Khalil, Jones, Babar, Jan, Zafar, and Alhussain (2019) reviewed studies on speech emotion recognition based on deep learning techniques. Khalil et al. (2019) identified deep learning techniques, such as the deep Boltzmann machine, recurrent neural networks, recursive neural networks, deep belief networks, convolutional neural networks, and auto encoders, and suggested that they offer easy model training and the efficiency in terms of shared weights. Liu and Wang (2018) focused on studies on gesture recognition in their review research. Liu and Wang (2018) covered gesture recognition technologies and algorithms and discussed how gesture recognition can be applied to human-robot collaborative manufacturing. Recognition technologies have great potential to assist learning and instruction. However, based on an analysis of the review studies referenced above, it can be concluded that they focused on one specific recognition technology only, and none of them considered recognition technologies in an educational context. To address this gap in the field, the present review study was carried out. In this review study, we identified different recognition technologies, described their recognition process, and explained their applications to learning and instruction with corresponding examples. Furthermore, schemes of the reviewed studies (e.g., participants, data collection, or research design) were also covered and the advantages and disadvantages of recognition technologies were discussed. This information may be useful for educators and researchers, especially those with a non-technical background, in terms of providing a greater understanding of recognition technologies and their recognition processes, applications, and schemes, as well as their advantages and disadvantages. Furthermore, this review study may provide ideas for future research on and development of the use of recognition technologies in the field of education.

## 2. Methodology

We conducted a systematic review in this study (Grant & Booth, 2009). The methodology was based on general recommendations proposed in related review studies (Shadiev & Sintawati, 2020a; Shadiev & Yang, 2020c). The methodology included three steps: a search, expert meetings, and a content analysis (Shadiev et al., 2014). First, we searched related studies on recognition technologies and their applications intended to assist learning and instruction (Figure 1). We searched such studies using the Web of Science database because Caseiro and Santos (2018) suggested that "it is one of the most extensive, popular and relevant research databases for the academic community" (p. 8). Our search was carried out in January 2020 and included such keywords as recognition, technology, learning, teaching, and instruction in different combinations because they were frequently used in earlier review studies (Shadiev et al., 2014). A total of 2,465 articles were returned by the search. After that, we used the following criteria for the purpose of narrowing down the selection of the research articles: studies published (1) within a last decade, between 2009 and 2019, (2) those focused on recognition technology applied to learning and instruction, (3) studies in English, and (4) studies appearing as a full text article in Social Science Citation Index (SSCI) journals related to educational technology. Duman, Orhon and Gedik (2015) argued that "SSCI journals adopt stringent criteria in reviewing articles. These articles are generally regarded as having higher impacts in the field" (p. 200). A total of seventy-two papers were selected and reviewed, and eighteen different recognition technologies were identified (see Appendix I).

Two researchers were involved in article selection process. First, they independently examined all articles against the above-mentioned criteria and removed all duplicate articles. After that, differences in the screening processes were identified and discussed. Finally, a consensus was reached among the researchers. We measured interrater reliability, which was high, with correlations of .95 between the researchers.

Second, the expert meeting was arranged following the general recommendations of Voogt and Roblin (2010). Eight experts were invited to the meeting. Each expert had a PhD degree and research expertise in educational technologies. The experts were asked to review eighteen different recognition technologies (see Appendix I) in order to discuss their theoretical and practical value as well as their features and functions. Based on the expert meeting, the following dimensions were derived: (1) recognition process, (2) application, (3) scheme, and (4) advantages and disadvantages.

Finally, a content analysis of seventy-two papers was carried out. It was based on open coding, and two researchers were involved in the process (Shadiev et al., 2014; Shadiev & Sintawati, 2020a; Shadiev & Yang, 2020c). The researchers carried out the coding process independently. They read the article and coded the data of interest with respect to the derived dimensions, including the recognition process, application of the recognition technology, scheme, and advantages and disadvantages of the given recognition technology. After that, the researchers finalized the coding phase through resolving any disagreements by re-examining the papers.

In the following section, we report our results and describe them with examples from the representative papers in order to enhance the reader's comprehension.

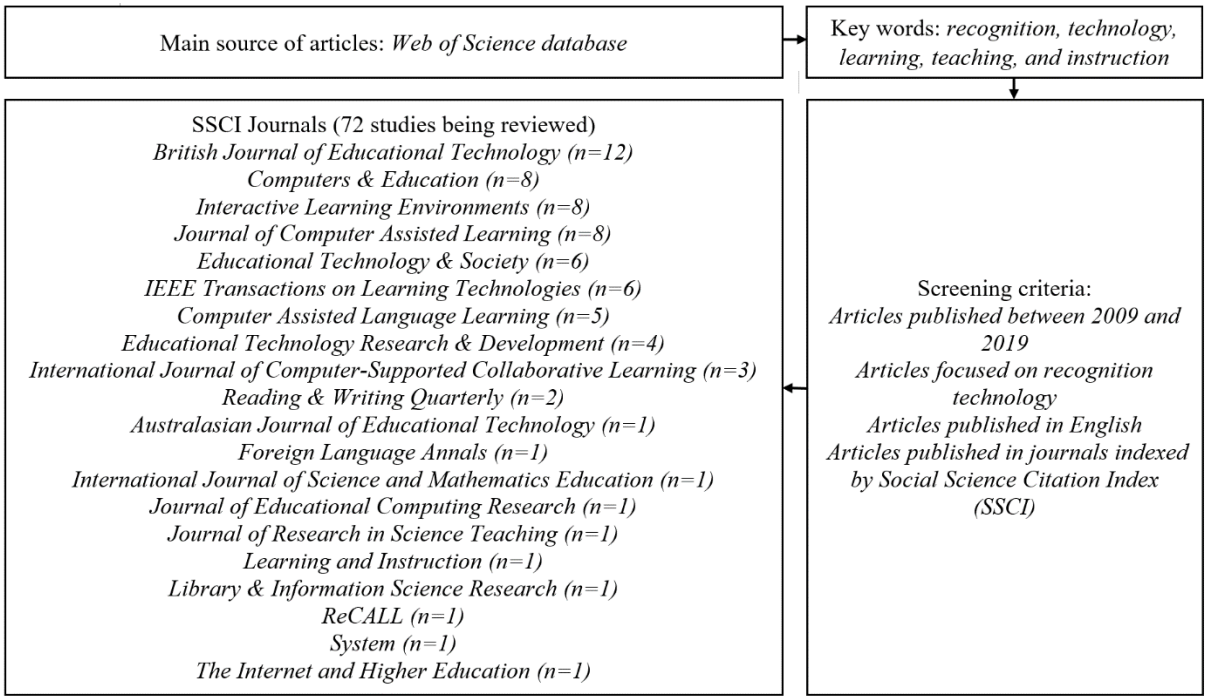


*Figure 1.* Search of articles

## 3. Results

### 3.1. Recognition technology and process

Figure 2 demonstrates the number of articles published by year. The highest number of articles was published in the last four years (i.e., in 2016 ($n$ = 11), 2017 ($n$ = 11), 2018 ($n$ = 13) and 2019 ($n$ = 10)), whereas the lowest number was published in 2009 ($n$ = 1). The publication trendline demonstrates that the trend in publishing articles is increasing.

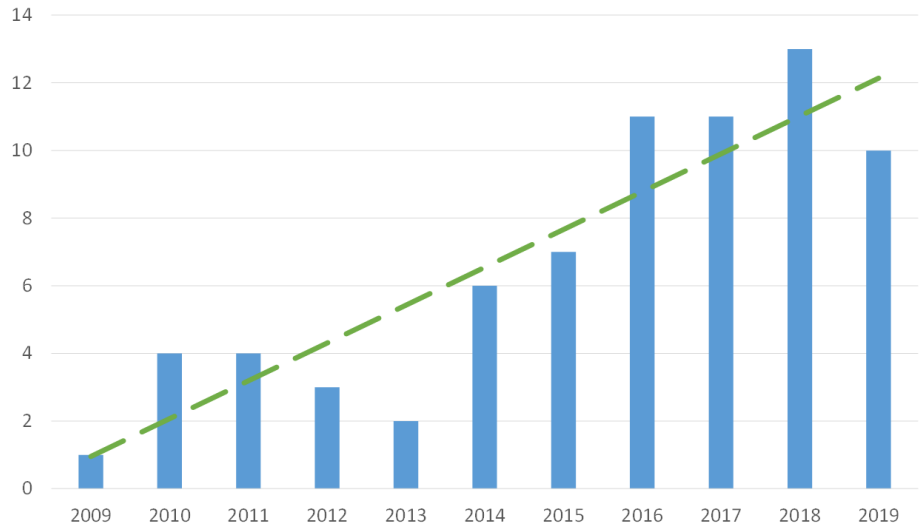


*Figure 2.* Published articles and publication trendline (broken green line)

Eighteen different recognition technologies were identified (see Figure 3). The most frequently used recognition technologies were speech recognition ($n$ = 16) and touch recognition ($n$ = 16) whereas the least frequently used were attention recognition ($n$ = 1), biometric information recognition ($n$ = 1), face recognition ($n$ = 1), hand

movement recognition ($n$ = 1), math expression recognition ($n$ = 1), physical activity recognition ($n$ = 1), sign recognition ($n$ = 1) and text-to-speech recognition ($n$ = 1).

All identified recognition technologies could be grouped into six categories based on the type of input data in the recognition processes: audio, non-text image, text image, moving image, physiological signals, and other (see Appendix I). In the following, we list these technologies and explain them in relation to the recognition processes.

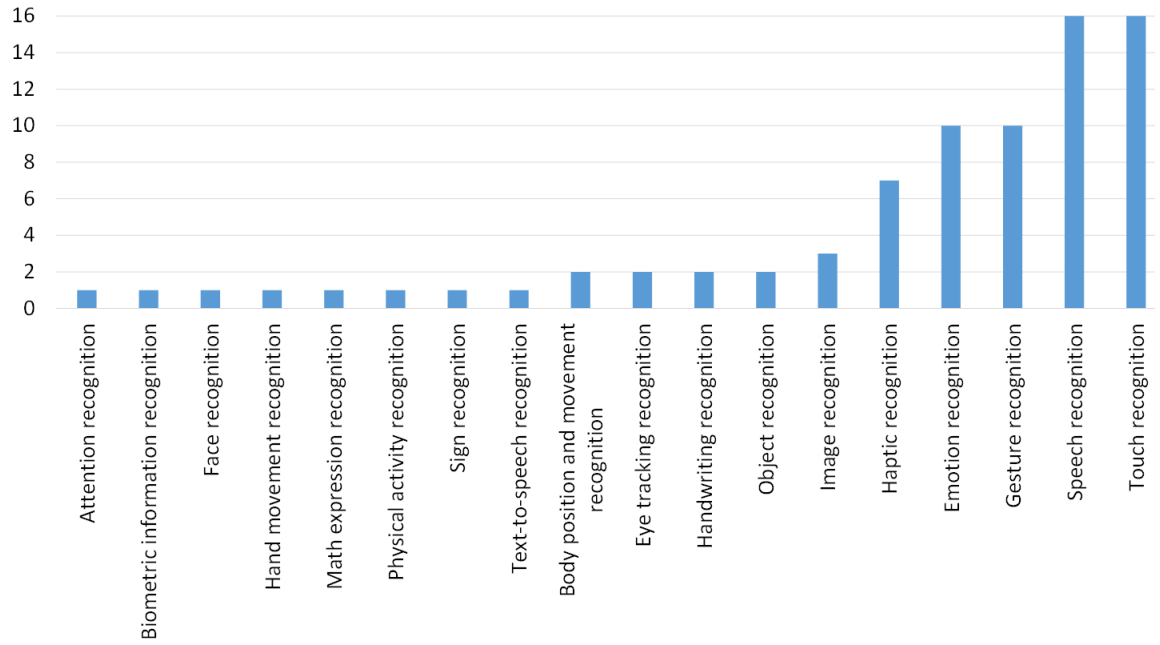


*Figure 3.* Recognition technologies

**Audio.** This category includes three types of recognition technologies based on audio: biometric information, speech, and emotion recognition technologies. Biometric information recognition technology records a student's voice using a microphone and then infers deception from stress based on a biometric voice stress analysis, i.e., the non-verbal, low-frequency content of the voice conveys information about the physiological and psychological state of the student (Kaklauskas et al., 2010). Similarly, emotion recognition technology receives voice input through a microphone and identifies the emotional states of students based on voice intonation (Bahreini, Nadolski & Westera, 2016). Speech-to-text recognition technology synchronously transcribes text streams from speech input that are simultaneously shown to students on a whiteboard or computer screens during lectures (Shadiev et al., 2014).

**Non-Text image.** Recognition technologies in this category are based on non-text images and include five types: face, image, sign, object, and emotion recognition technologies. For example, facial recognition technology captures the human face using a camera and compares the captured image with data stored in a database (Yueh et al., 2014). Image recognition technology detects marker labels or graphics through a camera, and then, relevant virtual information is displayed on the device when a target is recognized (Hwang et al., 2018). Sign recognition technology converts user input sign language into language text (Ditcharoen, Naruedomkul, & Cercone, 2010).

**Text image.** In this category, recognition technologies are based on text images and include handwriting and math expression recognition technologies. Handwriting recognition technology supports handwritten input through tablets, detects the results of student writing, and provides feedback (Yanikoglu et al., 2017). Math expression recognition technology recognizes written math symbols, converts them into formats that a computer can understand, and then displays the recognized math expressions on a computer screen (Pacheco-Venegas et al., 2015).

**Moving image.** This category consists of four recognition technologies that are based on moving images, i.e., body position and movement, gesture, eye tracking, and hand movement recognition technologies. Body position and movement recognition technology detects postures or movements made by students using Kinect, determines whether these postures or movements meet the requirements, and provides feedback (Wang et al., 2019). Gesture recognition technology captures live images of a user through a camera and recognizes user

gestures from the live images through computer vision technology intended to observe gesture interactions (Yang & Liao, 2014). Hand movement recognition technology records and processes a student's hand movement and responses to the movement through a Kinect sensor (Wolski & Jagodziński, 2019). Eye-tracking technology measures eye positions and eye movement to reveal where visual sensory information is being acquired and how much attention a student is paying during the learning process (Huang et al., 2015; Wiebe, Minogue, Jones, Cowley & Krebs, 2009).

**Physiological signals.** Three types of recognition technologies based on physiological signals are included in this category: attention, emotion, and physical activity. Attention recognition technology monitors the brainwaves of a student's mental activity, registers electrical impulses generated by the brain, and outputs a value representing the current state of attention (Chen et al., 2017). Emotion recognition technology senses a student's pulse signals through an ear sensor, analyzes variability in heart rate patterns, interprets the data, and notifies a student about his/her status and changes in emotional states (Chen & Wang, 2011). Physical activity recognition technology detects a student's physical activity through the sensors in a smart phone and a smart wrist band. This technology collects data such as heart rate through the optical heart rate sensor in the smart wrist band, and then this data is recorded into a file (Lindberg et al., 2016).

**Others.** This category includes four recognition technologies: haptic, touch, emotion, and text-to-speech. Haptic recognition technology enables students to manipulate virtual objects through a haptic interface and then causes students to feel the force feedback (Magana & Balachandran, 2017). Touch recognition technology allows users to directly touch and interact with content without having to use an input device (Mercier et al., 2017). The technology senses the touch of a user through a camera and executes a respective command or sends respective feedback.

### 3.2. Application

The results show that recognition technologies were applied to educational contexts in three different ways: (1) as a tool or method to assist learning, (2) to analyze learning behavior, and (3) to identify the learning status of students (see Appendix II).

**A tool or method to assist learning.** In this dimension, recognition technologies are reviewed in terms of their applications as tools or methods to assist learning. Researchers have studied the application of speech recognition to speaking and pronunciation practice in foreign or second languages. For example, scholars developed an application with automatic speech recognition to improve language learners' speaking skills (Ahn & Lee, 2016; Chen et al., 2016; McCrocklin, 2016; Mroz, 2018; Van Doremalen et al., 2016), where the application provides immediate feedback regarding how correct the language learners' pronunciation is. Furthermore, language learners can obtain remedial exercises based on the type of mistakes they make.

Scholars have used visuo-haptic simulators to teach different subjects by combining tactile and visual perception, e.g., mechanics concepts in Neri et al. (2018) and Wiebe et al. (2009), physics concepts in Hamza-Lup et al. (2010), a dental curriculum in San Diego et al. (2012), or robot programming concepts in Howard, Park and Remy (2012). Learners were able to change the relevant parameters as well as apply and feel forces in the experiments. They exerted forces by manipulating a haptic stylus. Scholars have argued that haptic feedback is helpful to students to conceptualize abstract, difficult concepts (Magana & Balachandran, 2017).

Many researchers have studied touch recognition technology. This technology has been applied to assist learning of art (Hung et al., 2015), math (Chen et al., 2016; Hwang, Shadiev, Tseng & Huang, 2015; Mercier, Higgins & Da Costa, 2014), geography (Hung et al., 2014), and science (Davis et al., 2015; Horn et al., 2016). In most studies, a multi-touch tabletop technology was employed to facilitate collaboration among learners (Ardito et al., 2013; Bause et al., 2018; Mercier et al., 2014; Schmitt & Weinberger, 2019). Some scholars compared the learning effects of touch recognition technology with other approaches, e.g., pen and paper (Higgins, Mercier, Burd & Joyce-Gibbons, 2012; Mercier et al., 2017) or tablet PCs (Hwang et al., 2015). They found that a multi-touch tabletop had a greater learning effect than other approaches (Hwang et al., 2015). For example, students had better interactions and idea development under multi-touch conditions (Mercier et al., 2017).

Image recognition technology also received considerable attention in reviewed studies (Hwang et al., 2018; Liou, Yang, Chen & Tarng, 2017; Shadiev, Wu & Huang, 2020b). For example, a mobile learning system with an image recognition function was developed in Hwang et al. (2018) and Liou et al. (2017). It used image-based AR technology to support the learning of a local culture (Hwang et al., 2018) or science (Liou et al., 2017). A learner

had to direct the system to a learning target, and when it was recognized, the information related to the target was displayed on a mobile device to guide learners or provide additional learning content.

Handwriting recognition technology was applied to handwriting diagnoses and remedial instruction in Hsiao et al. (2015) and Yanikoglu et al. (2017). The technology could detect errors in handwriting and provide immediate feedback. A biometric, intelligent self-assessment system based on a voice stress analysis and a special algorithm was used in Kaklauskas et al. (2010) to assist students with self-assessment. Face recognition technology was used in Yueh et al. (2014) to develop a learner identification system in order to support real-time interaction in an international distance course. The facial recognition component of the system was used to recognize learners. Wolski and Jagodzinsk (2019) developed a virtual chemical laboratory based on hand movement recognition technology. It detected the hand movement of students through Kinect, where the students could simulate chemical experiment activities in a virtual laboratory, such as operating specific laboratory equipment.

Pacheco-Venegas et al. (2015) used an automatic evaluation system for mathematical expressions. The system could recognize mathematical symbols written by students, evaluate the math expression, and finally give feedback. Ditcharoen et al. (2010) developed a system that was able to translate from Thai sign language into Thai text to help deaf students learn the Thai language. van Laere et al. (2017) used bilingual text-to-speech technology to provide auditory at-home support for students such as those who speak a language other than the language being used for teaching. Text-to-speech technology was used to read the content on a screen, allowing students to obtain both visual and auditory information simultaneously.

**Analysis of learning behavior.** In this dimension, we reviewed recognition technologies that were used for the analysis of learning behavior. That is, recognition technologies detected learning behavior, analyzed it, and then provided corrective feedback and / or guided learners (Barmaki et al., 2018). Gesture recognition technology received a significant amount of attention in the reviewed studies. For example, this technology was employed to assist with learning and writing Chinese characters (Hao et al., 2010; Hong et al., 2017), to improve fine motor skills and recognition in children with autism spectrum disorder (Cai et al., 2018), pottery wheel-throwing (Glushkova & Manitsaris, 2018), and English as a foreign language (Yang & Liao, 2014). Students in these studies used gestures to interact with digital objects, e.g., students used a potter's wheel, and the system provided feedback on how well the throwing moves were. The reviewed studies demonstrated the positive effects of gesture-based interactions on learning.

Body position and movement recognition was the second most popular technology for analysis of learning behavior. In the reviewed studies, students used various postures or movements to interact with the system, and it could detect these postures or movements and give feedback to the learners. For example, students in Wang et al. (2019) learning English as a foreign language practiced newly learned vocabulary using specific postures or movements to demonstrate the meaning of specific words.

Physical activity recognition technology was used in Lindberg et al. (2016). They developed an exercise game in which wearable devices could recognize students' physical movements and improve their physical education environment. The sensors in smart phones and smart wrist bands could detect students' physical movements such as shaking, jumping, and spinning, as well as collect their heart rate data.

Eye-tracking technology is able to reveal which information students are paying attention to and how much attention they are giving that information. Eye-tracking technology was employed in Huang et al. (2015) because students were exposed to multimedia learning content including three different types of visual sensory information (i.e., a video of the instructor, slides of the lecture, and lecture captions).

**Identifying students' learning status.** In this dimension, recognition technologies were reviewed in terms of their ability to identify the learning status of students. Emotion recognition was the most frequently used technology in this dimension. This technology can be used to recognize the learning status of students based on facial expressions or vocal intonations and provides appropriate help and guidance based on a learner's emotional state (Bahreini et al., 2016; Chickerur et al., 2015; Lin et al., 2016). For example, scholars used emotion recognition technology to improve communication skills (Bahreini et al., 2016), to predict students' emotional states (Daouas et al., 2018), to reduce second-language speaking anxiety (Chen & Lee, 2011), to assess the effects of different types of multimedia materials on emotions associated with learning as well as learning performance (Chen & Wang, 2011), and to improve learning effects in students (Lin et al., 2016). Only one study employed attention recognition technology. Chen et al. (2017) developed an attention-aware system to assess students' attention levels. The system could identify high and low attention levels in students in an autonomous e-learning environment.

### 3.3. Scheme

This section is divided into participants, data collection methods, research design, and course/discipline sub-sections. In the following paragraphs, each of them is discussed individually.

**Participants.** Among the 72 studies found, 40.3% ($n$ = 29) of the studies involved students from universities, e.g., undergraduate students, graduate students, and doctoral students (see Appendix III). 29.2% ($n$ = 21) of the studies were conducted among primary school students, where participants in primary school, elementary school, and children below sixth grade were classified as primary school students. 13.9% ($n$ = 10) of these studies involved high school and middle school students, mainly including high school students, vocational school students, and middle (or secondary) school students. 9.7% ($n$ = 7) of the studies involved other participants, e.g., employees or experts. In 16.6% ($n$ = 12) of the studies, the participant information was not specified. That is, the participants did not wish to reveal their demographic information. 1.4% ($n$ = 1) included pre-school students.

**Data collection methods.** The researchers used questionnaires/scales (69.4%, $n$ = 50), tests (47.2%, $n$ = 34), interviews (34.7%, $n$ = 25), observation (16.7%, $n$ = 12), and focus groups (5.6%, $n$ = 4) to collect research data (see Appendix IV). In addition, some researchers used identification technology systems to collect data. For example, Chen and Wang (2011) used the emWave system to measure the emotional state of learners. Shadiev, Huang and Hwang (2017) used NeuroSky MindWave™ to measure student attention and degree of meditation.

**Research design.** Among the studies found, 48.6% ($n$ = 35) carried out experiments and set up groups for comparison, followed by not specified (34.7%, $n$ = 25), within-subjects designs (5.6%, $n$ = 4) and mixed-methods designs (4.2%, $n$ = 3) (see Appendix V). The least frequently designs were single group quasi-experiment (1.4%, $n$ = 1), single subject research designs (1.4%, $n$ = 1), design-based research approaches (1.4%, $n$ = 1), and interdisciplinary design approaches (1.4%, $n$ = 1).

**Course/Discipline.** We listed the courses/disciplines in which recognition technologies were used (see Appendix VI). The researchers applied recognition technologies to language learning (20.8%, $n$ = 15), mathematics (15.3%, $n$ = 11), science education (7%, $n$ = 5), cultural learning (5.6%, $n$ = 4), history (5.6%, $n$ = 4), and pre-school education (1.4%, $n$ = 1). In addition to the above, the researchers also applied recognition technologies to other courses/disciplines. For example, physics, chemistry, PE, art, writing, dental curricula, robot programming, etc.

### 3.4. Advantages

The advantages of recognition technology were reported in terms of the following categories: audio, non-text image, text image, moving image, physiological signals, enhanced human-computer interaction, helping teachers more accurately understand the learning status of students, increasing student motivation and engagement, and promoting understanding during the learning process.

**Audio.** Biometric information recognition was found to be beneficial to learning in the reviewed studies. For example, Kaklauskas et al. (2010) concluded that a self-assessment system was superior to traditional systems because of the use of a biometric voice stress analysis. Scholars also proved the advantages of speech recognition technology to improve language learning. For example, Ahn and Lee (2016) used automatic speech recognition to improve speaking in EFL learners, where the speech recognition technology gave feedback in response to the speech input of students, and ensuing interaction helped the students interact socially and in turn stimulated their language learning. In another study, Shadiev et al. (2017) provided texts generated by speech-to-text recognition to non-native English speakers in English lectures. The texts helped students understand the lecture content.

**Non-Text image.** In terms of advantages of face recognition technology, according to Yueh et al. (2014), this technology can be used "to improve the learning involvement, teaching effectiveness, and quality of interaction in the context of distance education" (p. 191). The application of face recognition technology resolved earlier difficulties that the students had experienced related to seeing and recognizing one another during online discussions.

According to scholars, image-based AR technology helps students learn in real-world contexts (Hwang et al., 2018; Liou et al., 2017; Shadiev et al., 2020b). This technology recognizes objects in the real world and then creates and displays virtual information on the screen, where, for example, image-based AR recognized learning

targets, and relevant questions are presented to guide students to observe targets and locate supplementary materials (Hwang et al., 2018).

Advantages of sign recognition technology were also reported. In a study by Ditcharoen et al. (2010), a Thai sign language to Thai machine translation system helped deaf students learn Thai and reduced their dependence on teachers. The system was conducive to improving the interest and motivation of these students, as well as their engagement in learning.

**Text image.** Two studies related to handwriting recognition revealed that this technology helped students' learn handwriting and gave learners feedback based on the handwriting recognition results (Hsiao et al., 2015; Yanikoglu et al., 2017). For example, Yanikoglu et al. (2017) found that handwriting recognition technology helped learners with handwriting exercises by detecting areas where the students needed to improve their handwriting, and Hsiao et al. (2015) demonstrated that handwriting recognition technology improved students' ability to write Chinese characters and understand the spatial structure of the characters. In terms of math expression recognition technology, according to Pacheco-Venegas et al. (2015), this tool facilitated user-computer interaction, where users could use their handwriting to interact with computers.

**Moving image.** Advantages of recognition technologies with respect to moving images were also reported in the reviewed articles. For example, scholars argued that gesture recognition can be used to achieve intuitive interaction. By combining augmented reality and computer vision, Yang and Liao (2014) developed a virtual English classroom in which students interacted with virtual objects through gestures. The scholars also suggested that hand movement recognition technology can be used to recognize gestures or hand movements made by students, which makes the software more interactive. Wolski and Jagodzinsk (2019) developed a virtual chemical laboratory based on a hand motion recognition system that achieved good results in a chemistry education context. In the virtual laboratory, students could simulate chemical experiments using gestures.

**Physiological signals.** Attention, emotion, and physical activity recognition technology based on physiological signals also has advantages. For example, in the learning process, attention recognition technology helps instructors evaluate students' attention levels; emotion recognition technology can be used to help teachers understand changes in students' emotions, and physical activity recognition technology can be used to recognize students' level of physical activity so teaching strategies can be adjusted based on levels and changes in the attention, emotions and physical activity levels of students in order to improve learning outcomes. For example, in Chen et al. (2017), the recognition system identified high and low attention levels in students in an autonomous e-learning environment and notified the instructor so that appropriate interventions could be made. The results of the reviewed studies showed that the learning process was more efficient when recognition technologies were introduced in the learning process (Chen et al., 2017; Chen & Wang, 2011; Lindberg et al., 2016).

**Enhanced human-computer interaction.** In the reviewed studies, students interacted with a system through recognition technologies and received feedback from the system. The results showed that human-computer interaction was carried out through voice (Ahn & Lee, 2016) or face (Yueh et al., 2014) interfaces. In addition, students interacted with objects in the real world (Hwang et al., 2018), digital objects in virtual environments (Wolski & Jagodzinsk, 2019; Yang & Liao, 2014), or created digital content (Hsiao et al., 2015; Pacheco-Venegas et al., 2015; Yanikoglu et al., 2017) by interacting with computers. This human-computer interaction was beneficial for improving various learning outcomes, e.g., language learning skills (Ahn & Lee, 2016; Yang & Liao, 2014), socialization and online learning (Yueh et al., 2014), procedural skills in science experiments (Wolski & Jagodzinsk, 2019), and character or mathematic formula writing skills (Hsiao et al., 2015; Pacheco-Venegas et al., 2015; Yanikoglu et al., 2017). For example, Ahn and Lee (2016) claimed that recognition technology provided feedback to student speech input, fostered the students to interact more and stimulated their learning. According to Hwang et al. (2018), image-based AR technology helped students learn in real-world contexts. Image-based AR recognized learning targets, where relevant questions were presented to guide students to observe targets and locate supplementary materials. In Yanikoglu et al. (2017), handwriting recognition technology helped with handwriting exercises by detecting areas where students needed to improve their handwriting. Hsiao et al. (2015) found that their system improved learners' ability to write characters and understand the spatial structure of the characters.

**Helping teachers more accurately understand the learning status of students.** Recognition technologies in this category could be used to help instructors determine the learning status of students with better accuracy. For example, recognition technologies were useful for instructors to evaluate attention levels in students (Chen et al., 2017) or understand changes in their emotions (Chen & Wang, 2011) while they were learning. This made it

possible for the teachers to adjust their teaching strategies to respond to changes in attention or emotion in an effort to improve student learning performance.

**Increase student motivation and engagement.** Scholars in the reviewed studies reported that applying recognition technologies to the learning process was useful in terms of increasing student motivation and engagement. For example, Shadiev, Hwang and Liu (2018) suggested that recognition technology provides feedback on student physical activity levels, where students can compare their levels to the goals set by the instructor. If physical activity level contributed to accomplishing a goal, then the students were motivated to continue the physical activity; otherwise, feedback warned them that more effort was needed. The results of the review showed that learning was more efficient and that the system was conducive to improving interest and motivation, as well as learning engagement (Ditcharoen et al., 2010; Lindberg et al., 2016; Shadiev et al., 2018).

**Promote students' understanding during learning.** Recognition technologies were also found to be useful in terms of promoting student understanding during the learning process. For example, Shadiev et al. (2017) argued that those students for whom the main communication language at an academic event was not their first language were not able to fully understand the communicated content at the event. In such cases, recognition technologies could be used to assist their understanding of communicated content. Shadiev et al. (2017) provided texts generated by speech-to-text recognition to non-native English speakers in English lectures. The texts helped the students understand the lecture content.

### 3.5. Disadvantages

There are some limitations to recognition technologies for instruction or learning that were reported in the reviewed articles. The main limitation is the accuracy of the recognition technology. For example, some scholars reported that the recognition process was not 100% accurate (Yanikoglu et al., 2017) and that there were miss-detections and false alarms in the recognition process (Hwang et al., 2018; Lindberg et al., 2016; Pacheco-Venegas et al., 2015; Yang & Liao, 2014). One reason for this issue was the fact that the recognition technology was not mature enough (Shadiev et al., 2017). Another reason was because the stability and correctness of recognition technology can be affected by the characteristics of the signals it receives; e.g., they are not clear enough or are mixed with other unrelated signals (Hwang et al., 2018). The problem associated with the accuracy of the recognition technology can lead to distraction (Yang & Liao, 2014) or inaccurate feedback from the system (Yanikoglu et al., 2017). Therefore, recognition accuracy needs to be improved (Hsiao et al., 2015).

To address these limitations, perhaps the use of high-quality recognition equipment can improve the accuracy of recognition technology. In addition, the development of technology is now very rapid, so some new technologies may be helpful with regard to improving the accuracy of recognition technology. Also, the correct use of recognition equipment can improve the accuracy of recognition technology. Furthermore, it is possible to use multi-recognition strategies as well as adjusting the calibration of devices by using feedback or biofeedback in order to improve accuracy rates.

Some of the scholars mentioned other limitations. For example, students need to learn a new interface for recognition technologies, e.g., a gesture interface instead of using a mouse and keyboard (Yang & Liao, 2014), or wearable recognition technologies were not easy to use; e.g., headsets that detect brain waves were difficult to keep on the heads of the participants because they could slip from their position (Shadiev et al., 2017). Thus, the ease of use of recognition equipment needs to be taken into consideration (Shadiev et al., 2017; Yang & Liao, 2014). The novelty effect, i.e., the short-term improvement in learning outcomes that comes from increased interest in a new technology, was also mentioned in the literature (Shadiev et al., 2018). One may argue that other disadvantages of recognition technologies such as cost, privacy, security, and ethical issues need to be reported. For example, some recognition technologies may be too expensive and therefore inaccessible to many educators, researchers, and students. However, according to our review, no results related to any of these issues were reported in the reviewed articles.

## 4. Discussion and conclusion

Recognition technologies have great potential in education (Shadiev et al., 2014). However, to the best of our knowledge, the reviewed studies related to one specific recognition technology only. Therefore, there was a need for a review study that summarized all existing recognition technologies and their applications to the educational process. This study was conducted to address this gap in the literature. We reviewed seventy-two articles, and

eighteen recognition technologies were identified. The following main aspects of such technologies were discussed with respect to identified technologies: (1) the recognition process, (2) applications of recognition technologies, (3) the scheme applied in the reviewed studies, such as the participants, data collection, and research design, and (4) the advantages and disadvantages of recognition technologies. Because applying recognition technologies in education is a very promising research direction, our review study may be useful to educators and researchers, especially for those with non-technical backgrounds. For example, our results can help educators and researchers understand what recognition technologies currently exist, the nature of the recognition process, the schemes and applications of recognition technologies to the field of education, and their advantages and disadvantages. Therefore, our results may be useful for educators and researchers to learn more about recognition technologies and their applications to education and may provide ideas for future research on and development of the use of recognition technologies in the field of education.

### 4.1. The recognition process in recognition technologies

Educators and researchers can use various recognition technologies for learning and instruction (Shadiev et al., 2014). We identified eighteen recognition technologies and grouped them into audio, non-text image, text image, moving image, physiological signals, and other categories based on the input. Audio recognition technologies receive input from the voice of the student or instructor (e.g., a student pronounces a word). Therefore, this technology can be used by educators to design learning activities intended to facilitate student pronunciation in language classrooms (Ahn et al., 2016; van Doremalen et al., 2016). Non-text image recognition technologies work based on various non-text images (e.g., objects in a classroom or student faces). For example, facial recognition technology can be used to support affective learning by identifying student emotions and providing relevant feedback (Bahreini et al., 2016), e.g., cheering up a student if the emotion is not positive. Image recognition technology can be used during the learning process to capture and recognize target objects around students and to augment with additional learning content (Hwang et al., 2018; Liou et al., 2017). Text image recognition technologies capture text features in images (e.g., mathematical formulas, letters or characters). This technology can be useful in assisting students with improving their handwriting or drawing skills (Hsiao et al., 2015; Pacheco-Venegas et al., 2015). Moving image technologies constitute various input data such as gestures, e.g., a student draws a circle to activate a context menu (Holmes et al., 2018). Educators may consider this type to recognition technology to facilitate embodied learning (Barmaki et al., 2018; Yang & Liao, 2014). Physiological signals such as the electrical activity in the brain or changes in heart rate can be also measured by recognition technologies. Educators may consider them to measure learning attention, emotions, or physical activity, for example, when student learning attention decreases (Chen et al., 2017), or students become less physically active (Lindberg et al., 2016), educators may want to intervene in order to increase attention or physical activity. Other input data such as haptic (e.g., signals generated through bodily movement by means of physical contact with devices such as joysticks and other haptics) and touch (e.g., a student touches a screen on a device) can be received and processed using recognition technologies as well. Haptic recognition technologies can be applied to science education, where students have tactile experience and feedback that creates a sense of touch through vibrations, motion, or other forces, e.g., learning about mechanics (Neri et al., 2018; Wiebe et al., 2009). Educators may consider incorporating touch recognition technology for students to learn how to interact and collaborate during problem solving tasks (Higgins et al., 2012; Mercier et al., 2017).

In addition, the review results showed that the recognition methods used for the same recognition technologies can be different. For example, for gesture recognition, either Kinect (Lai et al., 2018) or the Leap Motion Controller (Cai et al., 2018) can be used. For emotion recognition, heart rate data (Chen & Wang, 2011), facial expressions, vocal intonations (Bahreini et al., 2016), or galvanic skin response (Lin et al., 2016) can be used. Therefore, educators and researchers may select appropriate recognition technologies to use for their teaching and learning processes based on their input. Furthermore, they may consider measuring the same research variable using more than one recognition technology. In this case, the reliability and validity of measurements will be stronger.

Another issue to consider is the applicability of the research results to other educational settings. That is, scholars have reported results that are quite encouraging on most occasions, but how their approach to using specific recognition technologies will fit different educational settings in terms of age, gender, culture, location, etc. requires further exploration in future studies. Educators and researchers may also want to think about the datasets used in their research. In fact, they are all different in most studies because of the lack of a standard database. For this reason, it is difficult to make any comparison of the results with other similar methodologies under the same conditions. Therefore, a standard dataset that reflects global probability is a sorely needed resource for research in this field.

### 4.2. The application of recognition technologies

Our results suggest that educators and researchers may want to consider applying recognition technologies to educational contexts with respect to three categories: as a tool or method to assist learning, to analyze learning behavior, or to identify the learning status of students. In terms of the first category, recognition technologies can be applied to support learning and instruction. The results of the majority of the studies in this review showed numerous educational benefits related to the use of recognition technologies. However, it should be noted that there are obvious tendencies in the use of recognition technologies in terms of the features used to assist specific learning activities and subjects. For example, speech-based recognition technologies were applied mostly to language learning contexts, e.g., to practice speaking skills in foreign or second languages (Ahn & Lee, 2016; Mroz, 2018; Shadiev et al., 2014; van Doremalen et al., 2016), and technologies based on haptic input were used mostly in science education, e.g., students learned concepts related to the field of mechanics by manipulating virtual objects with tactile-sensory tools (Neri et al., 2018; Hamza-Lup et al., 2010; Schönborn et al., 2011). Touch recognition technology was applied to support collaborative learning activities among students in most studies because it features large multi-touch operated desktops (Higgins et al., 2012; Hwang et al., 2015; Mercier et al., 2017). It was also noted that students used touch recognition technology to learn mathematics or history in most of the reviewed studies involving this type of technology.

In terms of the second category, recognition technologies can be used to analyze the learning behavior of students. This will help instructors determine the learning behavior of students and apply appropriate teaching strategies in order to improve learning outcomes. Scholars in most of the studies agreed that recognition technologies are valid, reliable tools for analyzing learning behavior. Educators and researchers should however be aware that a tendency was observed indicating that recognition technologies in this category were used for analyzing student body movements such as handwriting, gestures, or postures. Physical education and language learning were the two subjects in which recognition technologies for learning behavior were employed the most (Hao et al., 2010; Hong et al., 2017; Lindberg et al., 2016).

In terms of the third category, scholars in the reviewed studies used recognition technologies to identify the learning status of students, which is important information. If the learning status is not appropriate, i.e., students are inattentive, anxious, or bored, then the instructors need to intervene. A tendency was observed toward an increase in the number of studies focusing on detecting student emotions and attention levels. The data for emotion recognition generally comprised variations in student facial expressions, voice, and heart rate (Bahreini et al., 2016; Chen et al., 2017; Chickerur et al., 2015).

Our results showed several tendencies in the usage of recognition technologies with respect to the three categories discussed above. This can be explained by the capacity and features of recognition technologies to detect and process inputted information. For example, speech recognition technology receives input from the voice, and therefore, it was employed in a context language acquisition context, specifically to improve speaking skills. Nevertheless, there are other educational contexts (i.e., not following such tendencies) to which recognition technologies can be applied, and educators and researchers may want to think about innovative ways to do so. For example, speech recognition technology can be used in different areas other than the language learning domain. One possible application is using speech recognition technology for transcription of a lecturer's spoken content during lectures. This approach could be applied to any domain knowledge rather than only to language learning. Transcripts can be used by students during lectures to take notes, to enhance their understanding of lecture content when some important information is missed or misheard, and after lectures to complete homework assignments.

Another issue is that the research in this field is relatively thin. That is, in one decade, there were only seventy-two studies published. Therefore, more research should be done, especially with respect to those under-used recognition technologies in educational contexts such as attention recognition, biometric information recognition, and math expression recognition.

### 4.3. Schemes of the reviewed studies, such as the participants, data collection, and research design

In the reviewed studies, we found that the largest proportion of participants comprised university and primary school students. This finding suggests that educators and scholars in most of the reviewed studies used recognition technologies to support either higher or primary education. This finding demonstrates that scholars did not involve other subjects (e.g., middle/secondary school students, in-service teachers, or participants involved in on-the-job training). We also found that, in almost one fifth of the studies, the participants'

demographic information was not shown. Therefore, it is suggested that, in the future, studies can be carried out in educational settings involving these under-researched groups of learners. This would help enrich the knowledge base related to such groups. We also suggest that future studies provide the participants' demographic information because it is important for the generalizability of the results, as well as for the internal and external validity of the cohort design.

Many of the researchers used questionnaires to collect the participants' views and perceptions, such as their degree of satisfaction with using the technology. Researchers also often used tests to measure learning outcomes. Questionnaires are very popular instruments to collect subjective data since the data can be collected easily from a large population. On the other hand, tests are very reliable instruments and are used for experimental studies to show the effectiveness of interventions. This finding suggests that other instruments were under-used by the researchers. Therefore, it is suggested that educators and researchers consider using other instruments as well.

With respect to research design, nearly half of the studies set up group comparisons. Some studies compared the learning outcomes of the experimental group with those of a control group. This indicates that the most common experimental method in the reviewed studies was a group comparison. One major concern with this finding is the lack of any comparative analysis with other benchmark works in the field (i.e., other similar recognition technologies). In future studies, researchers may want to consider carrying out such comparative analyses because they may show how good/bad the results are compared to those when other similar recognition technologies are used. Our findings also show that, in previous studies, other research designs were considered less. For example, a design-based research approach was employed in only one study (Yanikoglu et al., 2017). In addition, descriptive, explanatory, or correlational research designs were not used at all. Therefore, we suggest that, in the future, scholars may want to use other designs in their research as well. According to our results, the research design was not specified in twenty-three studies, so scholars need to explicitly report such important information in the future.

We found that research on recognition technologies can be divided into two main streams: development (i.e., some scholars developed and evaluated recognition technologies in their research) and application (i.e., some scholars applied existing recognition technologies to support learning and instruction). In the former, the researchers introduced the recognition technology system they developed, the framework and model used, and then evaluated the technology. For example, Bahreini et al. (2016) developed emotional recognition software and validated the recognition results. Daouas et al. (2018) proposed an emotionally intelligent e-learning system, introduced its architecture and the model used, and conducted a validation test. In the latter, the researchers used existing recognition technology tools and applied them to support learning and instruction. For example, Hao et al. (2010) and Hong et al. (2017) designed games based on gesture recognition to help learners learn Chinese. Mercier et al. (2017) studied cooperative learning with a multi-touch technology. Based on this finding, we suggest that various recognition technologies can be developed and tested based on different models and frameworks, or scholars may want to use existing recognition tools to support teaching and learning processes.

### 4.4. Advantages and disadvantages of recognition technologies

It is important for educators and researchers to know the advantages and disadvantages of recognition technologies and their applications to education. Such knowledge will enable stakeholders to use known advantages and to minimize known disadvantages in order for learning and instruction to be successful.

Researchers have demonstrated that most recognition tools have the potential to support the teaching and learning process. Nevertheless, the results of this review study demonstrated that recognition technology accuracy is a common problem in studies on recognition technologies (Hsiao et al., 2015; Lindberg et al., 2016; Pacheco-Venegas et al., 2015; Yang & Liao, 2014). Therefore, it is suggested that, in the future, improving the accuracy of recognition technology remains an important issue.

Another issue in recognition technology-related research arises from the fact that some technologies are intrusive or uncomfortable to use. For example, some students had a hard time keeping headsets used in recognition technology on their heads (Shadiev et al., 2017). This issue implies that it will be impossible to apply some of recognition technologies in real-life situations when the input data gas to be received by technologies during typical learning or working processes. Therefore, it should be noted that most of reviewed studies tested the effectiveness of recognition technology applications in laboratory settings, so how student learning outcomes can be improved in the real world is still not known. In this regard, we can say that application of recognition technologies to assist learning and instruction is still in the early stages of research. Therefore, more studies are

required, especially in real world settings where learning usually takes place (e.g., real classrooms, schools, communities, or homes), and how to make technologies less intrusive or make students less uncomfortable when using technologies needs to be considered in the future. Furthermore, such issues as the novelty effect, cost, privacy, security, and ethical issues while using recognition technologies need to be considered and reported in future studies.

### 4.5. Trends

The results demonstrated that trend toward publication is increasing. This suggests that recognition technologies are getting better in terms of their accuracy rates, ease of use, and usefulness for learning and instruction. As a result, scholars are beginning to pay more attention to such technologies and their applications to education. The results also showed that some technologies have received little attention in research compared to other technologies. For example, biometric information recognition or sign recognition were the least frequently used, and speech recognition or touch recognition were the most frequently used. This finding suggests that educators who are planning to apply recognition technologies and seek information about their desired applications to education may find very little about such recognition technologies such as biometric information recognition or sign recognition. Therefore, more research related to applications of such technologies in education is needed.

Future studies may consider the extension of recognition technologies with other research tools. For example, recognition technologies can be combined with artificial intelligence to facilitate learning. Artificial intelligence (AI) contains multiple fields and content, and recognition technologies will ultimately be designed and developed for the artificial intelligence educational field. Therefore, multi-sensory recognition schemes will be a trend to improve accuracy in the future. In addition, fueled by AI technology, recognition technology will become more accurate and smarter. In addition, the integration of virtual reality and recognition technology such as image-based AR is an important research issue worth investigating. In the future, researchers may also want to consider combining multiple recognition technologies to enhance learning. For example, it is possible to combine face and speech recognition in order to receive image and audio input from recognition systems that will improve accuracy rates and also make such systems more reliable. Finally, future recognition technologies can identify intrinsic mental and psychological activities, such as imagination, attention, and creativity, more accurately, and perhaps, future recognition machines can be trained to recognize targets more accurately.

### 4.6. Research limitations

Several limitations of this study need to be acknowledged and addressed in the future. We reviewed eighteen different recognition technologies in the application domains but not the technologies themselves. Therefore, scholars may wish to address this limitation in future studies.

In this review study, we only included articles that our search yielded based on specific search terms. This approach could have excluded some articles on certain recognition technologies if their author/s did not use any of these terms (e.g., recognition). For example, our search yielded, and we reviewed, only three studies related to image-based AR (see Appendix I). However, one may argue that there are potentially many more related studies. Therefore, future studies may consider including other search terms in addition to those that were used in this study, e.g., future studies may use such search terms as “augmented reality” in order to identify more studies related to image-based AR.

Another limitation relates to certain aspects of recognition technology applications intended to assist learning and instruction, such as its efficiency, accuracy, and cost. It would be very useful for educators and researchers to know how efficient, accurate, and costly certain recognition technology is and the relationship between such aspects and technology application for effective and successful learning and instruction, e.g., how accurate a recognition technology is and how its accuracy impacts learning and instruction outcomes. Exploring these aspects of recognition technologies were out of scope of the present research and should be considered by scholars in future review studies.

## References

Ahn, T. Y., & Lee, S. M. (2016). User experience of a mobile speaking application with automatic speech recognition for EFL learning. *British Journal of Educational Technology*, *47*(4), 778-786.

Ardito, C., Lanzilotti, R., Costabile, M. F., & Desolda, G. (2013). Integrating traditional learning and games on large displays: An Experimental study. *Educational Technology & Society*, *16*(1), 44–56.

Bahreini, K., Nadolski, R., & Westera, W. (2016). Towards multimodal emotion recognition in e-learning environments. *Interactive Learning Environments, 24*(3), 590-605.

Barmaki, R., & Hughes, C. E. (2018). Embodiment analytics of practicing teachers in a virtual immersive environment. *Journal of Computer Assisted Learning*, *34*(4), 387-396.

Bause, I. M., Brich, I. R., Wesslein, A. K., & Hesse, F. W. (2018). Using technological functions on a multi-touch table and their affordances to counteract biases and foster collaborative problem solving. *International journal of computer-supported collaborative learning*, *13*(1), 7-33.

Cai, S., Zhu, G., Wu, Y. T., Liu, E., & Hu, X. (2018). A Case study of gesture-based games in enhancing the fine motor skills and recognition of children with autism. *Interactive Learning Environments*, *26*(8), 1039-1052.

Calvo, M. G., & Nummenmaa, L. (2016). Perceptual and affective mechanisms in facial expression recognition: An Integrative review. *Cognition and Emotion*, *30*(6), 1081-1106.

Caseiro, N., & Santos, D. (Eds.). (2018). *Smart specialization strategies and the role of entrepreneurial universities*. Hershey, PA: IGI Global.

Chen, C. M., & Lee, T. H. (2011). Emotion recognition and communication for reducing second-language speaking anxiety in a web-based one-to-one synchronous learning environment. *British Journal of Educational Technology, 42*(3), 417-440.

Chen, C. M., & Wang, H. P. (2011). Using emotion recognition technology to assess the effects of different multimedia materials on learning emotion and performance. *Library & Information Science Research, 33*(3), 244-255.

Chen, C. M., Wang, J. Y., & Yu, C. M. (2017). Assessing the attention levels of students by using a novel attention aware system based on brainwave signals. *British Journal of Educational Technology, 48*(2), 348-369.

Chickerur, S., & Joshi, K. (2015). 3D face model dataset: Automatic detection of facial expressions and emotions for educational environments. *British Journal of Educational Technology*, *46*(5), 1028-1037.

Daouas, T., & Lejmi, H. (2018). Emotions recognition in an intelligent elearning environment. *Interactive Learning Environments*, *26*(8), 991-1009.

Davis, P., Horn, M., Block, F., Phillips, B., Evans, E. M., Diamond, J., & Shen, C. (2015). "Whoa! We're going deep in the trees!": Patterns of collaboration around an interactive information visualization exhibit. *International Journal of Computer-Supported Collaborative Learning, 10*(1), 53-76.

Ditcharoen, N., Naruedomkul, K., & Cercone, N. (2010). SignMT: An Alternative language learning tool. *Computers & Education*, *55*(1), 118-130.

Dong, G., Wang, N., & Kuang, G. (2014). Sparse representation of monogenic signal: With application to target recognition in SAR images. *IEEE signal processing letters*, *21*(8), 952-956.

Duman, G., Orhon, G., & Gedik, N. (2015). Research trends in mobile assisted language learning from 2000 to 2012. *ReCALL*, *27*(2), 197-216.

Glushkova, A., & Manitsaris, S. (2018). Gesture recognition and sensorimotor learning-by-doing of motor skills in manual professions: A Case study in the wheel-throwing art of pottery. *Journal of Computer Assisted Learning*, *34*(1), 20-31.

Ghosh, D., Dube, T., & Shivaprasad, A. (2010). Script recognition—A Review. *IEEE Transactions on pattern analysis and machine intelligence*, *32*(12), 2142-2161.

Grant, M. J., & Booth, A. (2009). A Typology of reviews: An Analysis of 14 review types and associated methodologies. *Health Information and Libraries Journal, 26*(2), 91–108.

Hamza-Lup, F. G., & Stanescu, I. A. (2010). The Haptic paradigm in education: Challenges and case studies. *The Internet and Higher Education*, *13*(1-2), 78-81.

Hao, Y., Hong, J. C., Jong, J. T., Hwang, M. Y., Su, C. Y., & Yang, J. S. (2010). Non-native Chinese language learners' attitudes towards online vision-based motion games. *British Journal of Educational Technology*, *41*(6), 1043-1053.

Higgins, S., Mercier, E., Burd, L., & Joyce-Gibbons, A. (2012). Multi-touch tables and collaborative learning. *British Journal of Educational Technology, 43*(6), 1041-1054.

Horn, M. S., Phillips, B. C., Evans, E. M., Block, F., Diamond, J., & Shen, C. (2016). Visualizing biological data in museums: Visitor learning with an interactive tree of life exhibit. *Journal of Research in Science Teaching*, *53*(6), 895-918.

Howard, A. M., Park, C. H., & Remy, S. (2012). Using haptic and auditory interaction tools to engage students with visual impairments in robot programming activities. *IEEE transactions on learning technologies*, *5*(1), 87-95.

Hsiao, H. S., Chang, C. S., Chen, C. J., Wu, C. H., & Lin, C. Y. (2015). The Influence of Chinese character handwriting diagnosis and remedial instruction system on learners of Chinese as a foreign language. *Computer Assisted Language Learning*, *28*(4), 306-324.

Hwang, G. J., Chang, S. C., Chen, P. Y., & Chen, X. Y. (2018). Effects of integrating an active learning-promoting mechanism into location-based real-world learning environments on students' learning performances and behaviors. *Educational Technology Research & Development*, *66*(2), 451-474.

Hwang, W. Y., Shadiev, R., Tseng, C. W., & Huang, Y. M. (2015). Exploring effects of multi-touch tabletop on collaborative fraction learning and the relationship of learning behavior and interaction with learning achievement. *Educational Technology & Society, 18*(4), 459-473.

Kaklauskas, A., Zavadskas, E. K., Pruskus, V., Vlasenko, A., Seniut, M., Kaklauskas, G., Matuliauskaite, A., & Gribniak, V. (2010). Biometric and intelligent self-assessment of student progress system. *Computers & Education*, *55*(2), 821-833.

Khalil, R. A., Jones, E., Babar, M. I., Jan, T., Zafar, M. H., & Alhussain, T. (2019). Speech emotion recognition using deep learning techniques: A Review. *IEEE Access*, *7*, 117327-117345.

Lin, H. C. K., Su, S. H., Chao, C. J., Hsieh, C. Y., & Tsai, S. C. (2016). Construction of multi-mode affective learning system: Taking affective design as an example. *Educational Technology & Society*, *19*(2), 132-147.

Lindberg, R., Seo, J., & Laine, T. H. (2016). Enhancing physical education with exergames and wearable technology. *IEEE Transactions on Learning Technologies*, *9*(4), 328-341.

Liou, H. H., Yang, S. J., Chen, S. Y., & Tarng, W. (2017). The Influences of the 2D image-based augmented reality and virtual reality on student learning. *Journal of Educational Technology & Society*, *20*(3), 110-121.

Liu, H., & Wang, L. (2018). Gesture recognition for human-robot collaboration: A Review. *International Journal of Industrial Ergonomics*, *68*, 355-367.

Magana, A. J., & Balachandran, S. (2017). Unpacking students conceptualizations through haptic feedback. *Journal of Computer Assisted Learning*, 33(5), 513-531.

McCrocklin, S. M. (2016). Pronunciation learner autonomy: The Potential of automatic speech recognition. *System*, *57*, 25-42.

Mercier, E. M., Higgins, S. E., & Da Costa, L. (2014). Different leaders: Emergent organizational and intellectual leadership in children's collaborative learning groups. *International Journal of Computer-Supported Collaborative Learning, 9*(4), 397-432.

Mercier, E., Vourloumi, G., & Higgins, S. (2017). Student interactions and the development of ideas in multi-touch and paper-based collaborative mathematical problem solving. *British Journal of Educational Technology*, *48*(1), 162-175.

Mroz, A. (2018). Seeing how people hear you: French learners experiencing intelligibility through automatic speech recognition. *Foreign Language Annals*, *51*(3), 617-637.

Neri, L., Noguez, J., Robledo-Rella, V., Escobar-Castillejos, D., & Gonzalez-Nucamendi, A. (2018). Teaching of classical mechanics concepts using visuo-haptic simulators. *Educational Technology & Society, 21*(2), 85–97.

Pacheco-Venegas, N. D., López, G., & Andrade-Aréchiga, M. (2015). Conceptualization, development and implementation of a web-based system for automatic evaluation of mathematical expressions. *Computers & Education*, *88*, 15-28.

San Diego, J. P., Cox, M. J., Quinn, B. F., Newton, J. T., Banerjee, A., & Woolford, M. (2012). Researching haptics in higher education: The Complexity of developing haptics virtual learning systems and evaluating its impact on students' learning. *Computers & Education*, *59*(1), 156-166.

Shadiev, R., Hwang, W. Y., Chen, N. S., & Huang, Y. M. (2014). Review of Speech-to-Text Recognition technology for enhancing learning. *Educational Technology & Society, 17*(4), 65–84.

Shadiev, R., Hwang, W. Y., & Liu, T. Y. (2018). A Study of the use of wearable devices for healthy and enjoyable English as a foreign language learning in authentic contexts. *Educational Technology & Society, 21*(4), 217–231.

Shadiev, R., Huang, Y. M., & Hwang, J. P. (2017). Investigating the effectiveness of speech-to-text recognition application on learning performance, attention, and meditation. *Educational Technology Research and Development, 65*(5), 1239-1261.

Shadiev, R., & Sintawati, W. (2020a). A Review of research on intercultural learning supported by technology. *Educational Research Review*, 100338. doi:10.1016/j.edurev.2020.100338

Shadiev, R., Wu, T. T., & Huang, Y. M. (2020b). Using image-to-text recognition technology to facilitate vocabulary acquisition in authentic contexts. *ReCALL, 32*(2), 195-212.

Shadiev, R., & Yang, M. (2020c). Review of studies on technology-enhanced language learning and teaching. *Sustainability*, *12*(2), 524. doi:10.3390/su12020524

Shih, F. Y. (2010). *Image processing and pattern recognition: Fundamentals and techniques*. Hoboken, NJ: John Wiley & Sons.

Schmitt, L. J., & Weinberger, A. (2019). Fourth graders' dyadic learning on multi-touch interfaces—Versatile effects of verbalization prompts. *Educational Technology Research and Development, 67*(3), 519-539.

van De Ville, D., Demesmaeker, R., & Preti, M. G. (2017). When Slepian meets Fiedler: Putting a focus on the graph spectrum. *IEEE Signal Processing Letters*, *24*(7), 1001-1004.

van Doremalen, J., Boves, L., Colpaert, J., Cucchiarini, C., & Strik, H. (2016). Evaluating automatic speech recognition-based language learning systems: A case study. *Computer Assisted Language Learning*, *29*(4), 833-851.

van Laere, E., & van Braak, J. (2017). Auditory support in linguistically diverse classrooms: Factors related to bilingual text-to-speech use. *Journal of Computer Assisted Learning*, *33*(1), 79-95.

Voogt, J., & Roblin, N. P. (2010). *21st century skills discussion paper.* Enschede, The Netherlands: University of Twente Press.

Wang, F., Hwang, W. Y., Li, Y. H., Chen, P. T., & Manabe, K. (2019). Collaborative kinesthetic EFL learning with collaborative total physical response. *Computer Assisted Language Learning*, *32*(7), 745-783.

Wiebe, E. N., Minogue, J., Jones, M. G., Cowley, J., & Krebs, D. (2009). Haptic feedback and students' learning about levers: Unraveling the effect of simulated touch. *Computers & Education*, *53*(3), 667-676.

Wolski, R., & Jagodziński, P. (2019). Virtual laboratory—Using a hand movement recognition system to improve the quality of chemical education. *British Journal of Educational Technology*, *50*(1), 218-231.

Yang, M. T., & Liao, W. C. (2014). Computer-assisted culture learning in an online augmented reality environment based on free-hand gesture interaction. *IEEE Transactions on Learning Technologies*, *7*(2), 107-117.

Yanikoglu, B., Gogus, A., & Inal, E. (2017). Use of handwriting recognition technologies in tablet-based learning modules for first grade education. *Educational Technology Research & Development*, *65*(5), 1369-1388.

Yueh, H. P., Lin, W., Liu, Y. L., Shoji, T., & Minoh, M. (2014). The Development of an interaction support system for international distance education. *IEEE Transactions on Learning Technologies*, *7*(2), 191-196.

## Appendix

All appendices are available from the following URL: https://yadi.sk/i/R-eiKEIfnEYUIw