

Article

An Improved Method for Online Teacher Training Courses for Bilingual Education Based on Cyber-Physical Systems

Ting-Hsuan Chien ¹, Yi-Lin Chen ¹, Jain-Shing Wu ^{1,*}, Cheng-Yan Siao ², Li-Ren Chien ²
and Rong-Guey Chang ²

¹ Department of Computer Science and Information Management, Providence University, Taichung City 43301, Taiwan; thchien0616@pu.edu.tw (T.-H.C.); ylchen3@pu.edu.tw (Y.-L.C.)

² Department of Computer Science and Information Engineering, National Chung Cheng University, Chiayi City 621301, Taiwan; hcyen103m@cs.ccu.edu.tw (C.-Y.S.); ccumouse@cs.ccu.edu.tw (L.-R.C.); rgchang@cs.ccu.edu.tw (R.-G.C.)

* Correspondence: jainshingwu1@pu.edu.tw; Tel.: +886-04-2632-8001 (ext. 18126)

Abstract: In recent years, bilingual education has become a critical index, and with its growth, internationalization has been significantly improved in the Republic of China (Taiwan). Therefore, countries worldwide are promoting all-English teaching, taking English as a medium of instruction and indicator of university education while moving toward an internationalized curriculum and teaching excellence. Finding the required number of bilingual teachers is also one of the keys to increasing the index. If information technology can be adopted to reduce the time and cost of bilingual teacher training, it will significantly improve the effectiveness of bilingual education promotion. However, the current traditional online training system only focuses on the delivery of one-way training content, and it cannot assist in judging the training effectiveness of the trainees. In this paper, we integrate the pose estimation technology into the online teacher training system to analyze the interactive content of the trainees during the training. We can assist in recording and interpreting the teaching demonstration process in classroom observation. We also verify the result by comparing our method's efficiency and judgment accuracy with the traditional way. The results show that our approach is more convenient and cost-efficient.

Keywords: distance learning; cyber-physical system; bilingual education; online judge system; English as a medium of instruction



Citation: Chien, T.-H.; Chen, Y.-L.; Wu, J.-S.; Siao, C.-Y.; Chien, L.-R.; Chang, R.-G. An Improved Method for Online Teacher Training Courses for Bilingual Education Based on Cyber-Physical Systems. *Appl. Sci.* **2022**, *12*, 2346. <https://doi.org/10.3390/app12052346>

Academic Editor: Alexander E. Hramov

Received: 30 December 2021

Accepted: 22 February 2022

Published: 23 February 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Many studies have tried to combine the two worlds of the Internet and the physical world in recent years. The cyber-physical systems (CPS) concept has also begun to emerge to integrate the interaction of computer systems with objects in the physical world. In this context, a technological impact on the learning environment has also continued to occur. New research directions have emerged in the education field, such as e-learning, distance learning, virtual learning environments, and mobile learning. E-learning is a broad term that usually includes all technologies and systems related to teaching and learning through electronic communication and digital technology. The evolution of e-learning has produced a series of new teaching materials using multimedia. The available resources in the teaching process have also become diversified in this case. The learning objects can start to obtain the help they need through methods including videos, images, animations, simulations, websites, and slides.

However, the current application of related virtual and real integration technology in education is mainly one way, simply transforming physical content teaching materials into virtual digital teaching materials for learning objects. Still, the physical interaction of the learning objects themselves is unresponsive. In terms of virtual content, bilingual teacher training includes interacting with students and innovative educational concepts,

such as constructing learning maps, learning history files, and precision education methods to design course content, so the focus of the feedback in the form of training results lies in the training. The process and content involve whether the bilingual teacher’s actual class hours and the range of the produced textbooks conform to the concept of bilingual teaching during the training. The existing traditional online training system is not suitable for training content that includes interaction with bilingual teacher training. Responding to the interactive content of the trainees of the virtual system with two-way virtual and real integration technology is the key to digital and remote bilingual teacher training.

Under the policy plan of promoting bilingual teaching and new southbound learning in Taiwan, the degree of internationalization of education has been correspondingly increased. This environmental factor has also significantly increased the learners’ demand for second language learning. The Ministry of Education is responding to the National Development Council’s goal of promoting the “Blueprint for the Development of a Bilingual National Policy in 2030”. The Ministry of Education’s strategy promotes teaching activation and life and stimulates learning motivation goals. The plan is to strengthen English listening and speaking practice in primary and secondary schools, establish an all-English teaching research center, and conduct all-English teaching teacher training courses. The Ministry of Education proposes that some fields or subjects in primary and secondary schools are taught in English and gradually implement English as a medium of instruction (EMI) in primary and secondary schools.

As for university higher education, EMI has become an indicator of the internationalization of current higher education courses and has also become an assessment program for academic performance and teaching excellence. However, promoting EMI is not just changing traditional classroom teaching based on English but also enabling EMI. The innovative learning model and student-centered teaching bring hope through the interaction between teachers and students. The students’ sense of participation and learning effect can be improved to achieve the results of bilingual education and learning, which is a situation for language and knowledge. Therefore, bilingual teachers’ training needs to be exchanged with the traditional education model.

In the face of the rapidly increasing demand for bilingual teachers, the Ministry of Education has also proposed related policies and measures, which are shown in Table 1. Using information technology and innovative educational methods to assist teachers who want to invest in bilingual education to conduct more efficient training can significantly reduce education costs and labor. By using online training, through synchronous or asynchronous training courses, there is no need to convene students worldwide, thus saving training costs. The venue will not limit exercise, and the training content can be updated in real time and at a low price. This makes full use of many online multimedia resources to enhance the richness of the course, and students can also decide their own learning time.

Table 1. Comparison of learning process and traditional assessment methods.

Target	Method
Accelerate teaching activation and life, stimulating learning motivation.	<ul style="list-style-type: none"> • Strengthen English listening and speaking practice in primary and secondary schools. • Promote some fields or subjects in primary and secondary schools to be taught in English. • Gradually implement English classes in primary and secondary schools to be taught in English. • Promote the teaching of practical English in professional groups in English in higher vocational colleges.

Table 1. Cont.

Target	Method
Expand English-speaking human resources to meet local needs.	<ul style="list-style-type: none"> Use resources such as international students to assist English teaching in primary and secondary schools, enhance the energy of English teaching resource centers in the local area, set up all-English teaching research centers, and conduct all-English teacher training courses.
Use technology to popularize individualized learning and develop unlimited possibilities.	<ul style="list-style-type: none"> Inventory and integrate central, local, and private websites and learning resources, use live broadcasts to learn together, strengthen digital learning, and use digital learning partners.
Promote the internationalization of the education system.	<ul style="list-style-type: none"> Expand international exchanges between schoolteachers and students (such as promoting international educational travel or Model United Nations). Promote the dual education system of colleges and universities. Add an international college or international degree program. Use the resources of the social education center (such as producing and broadcasting diversified English radio programs).
Loosen regulations, establishing flexible mechanisms and innovating learning models.	<ul style="list-style-type: none"> Accelerate the loosening of the “National Education Law,” the “Higher Secondary Education Law” (providing the legal source for setting up bilingual schools), and the “Private School Law” (opening foreign national schools to enroll students of their nationalities).

2. Related Work

2.1. Distance Learning

One of the early forms of distance learning was accomplished through correspondence courses in Europe. It became more and more popular [1] until the middle of this century through radios and TVs. As technology changes, the definition of distance learning has also changed. Videotaped lectures [2] have become the standard for universities and professional courses in the past two decades. Audiotapes and courses sent by mail have been used in correspondence courses to teach foreign languages for quite some time [3]. Nowadays, the Internet and compressed video have brought distance learning in a new direction, as distance learning can be carried out in real time.

As the Internet spreads across various industries, universities have also begun to use multiple pieces of software for online education [4–7]. As a result, distance learning has become increasingly popular, greatly expanding the teaching space [8] and allowing students to study in-depth inter-college university courses [9].

2.2. Cyber-Physical Integration Technology for Education Application

In the application of virtual-real integration technology in education, research proposed by Olympiou and Zacharia [10] involved making the students understand and experience the concepts related to light and color through the manipulation of physical objects and virtual objects. This study used clearly defined question sets and procedures to establish an agreement to evaluate students’ learning. By comparing objects that use traditional teaching methods with virtual-real integration methods, the results show that

the use of virtual-real integration methods provides students with a better understanding of the process.

Jones's research [11] involved a high degree of interaction with physical objects, though the interactive behavior between learners did not seem to be exploited. Regarding related technologies, the research only used virtual environments to perform virtual operations on objects and thus could not perfectly integrate information from the physical world for use in virtual environments. However, the research has a sufficient foundation for applying virtual and real integration systems in education. The research on integrating virtual and real education applications is good review material. On the other hand, the possibility of cooperation between students during the experiment is not discussed in the paper. It is related to virtual and physical objects, and there is no relevance in the tools provided by the author. To summarize the arguments of this research, teachers can identify a good student's actual operating state in the virtual-real integration system because the virtual-real integration system allows them to perform physical interactions such as manipulation, touch, vision, and movement of objects.

Lei and Man proposed an intelligent laboratory built through a virtual and real integration system [12]. The intelligent laboratory uses thermal comfort and safety as environmental adjustment factors. This intelligent laboratory was tested in a 150-student university engineering laboratory. The basis used in the research was to have an interactive algorithm with a temperature measurement sensor and an actuator that analyzed the current laboratory state to analyze the environment. This could be adjusted and operated appropriately in this intelligent laboratory.

After this research has determined the purpose of the virtual and real integration system, as the practical application of the research theory, students will face the development challenges of mechanical engineering and the handling of abnormal design control in the intelligent experiment. The experiment's plan is to control the temperature and humidity of the laboratory. For example, at a specific moment, the steam source is placed near the temperature sensor to simulate the high temperature and humidity changes in the environment. The students' solution must detect this situation and take action. For example, power is supplied to the fan's drive to provide sufficient ventilation and adjust the level in the environment.

2.3. Online Judge System

The online judge system allows users to send the completed code to the system for correctness verification according to the requirements of a given problem and to obtain the final result. The most commonly used result reference in the industry is Leetcode. As one of the applications of this type of system, it collects archaeological questions for software engineer interviewing and allows users to solve the questions on the website and use the results as a company reference.

In terms of learning and competitions, UVa Online Judge [13] or ACM online judge [14], as well as the more well-known ZeroJudge [15] in Taiwan, also provide many question banks and practice spaces for program-related learning and competitions. Other systems with different architectures, such as MetaOJ [16], have been developed for advanced usage, as well as many open-source online problem-solving techniques on GitHub, such as NOJ [17], DMOJ [18], and QDUOJ [19]. The online problem-solving approach mainly focuses on providing learners with a question bank and space for practice. Therefore, cooperating with these systems to guide students to participate in classroom learning and stimulate their follow-up autonomous use of similar system exercises is a critical point in the introduction of mixed learning in programming education.

3. Analysis

3.1. Teaching Video Analysis for Pose Estimation

We analyzed the teaching videos in the physical source and judged the trainees' activity levels during class and the state of interaction with the students according to each

image frame in the video. The teaching status and the analysis process were determined by the system from the video as shown in Figure 1. Different communication protocols processed the physical source signal of the teaching process in the training process through the splitter to obtain the encapsulated video data. Then, the audio and video streams of the video were obtained by unpacking according to the encapsulation format used. The audio stream will be used in other analyses. We used a video stream and imported it into the video analyzer for decoding and image pre-processing.

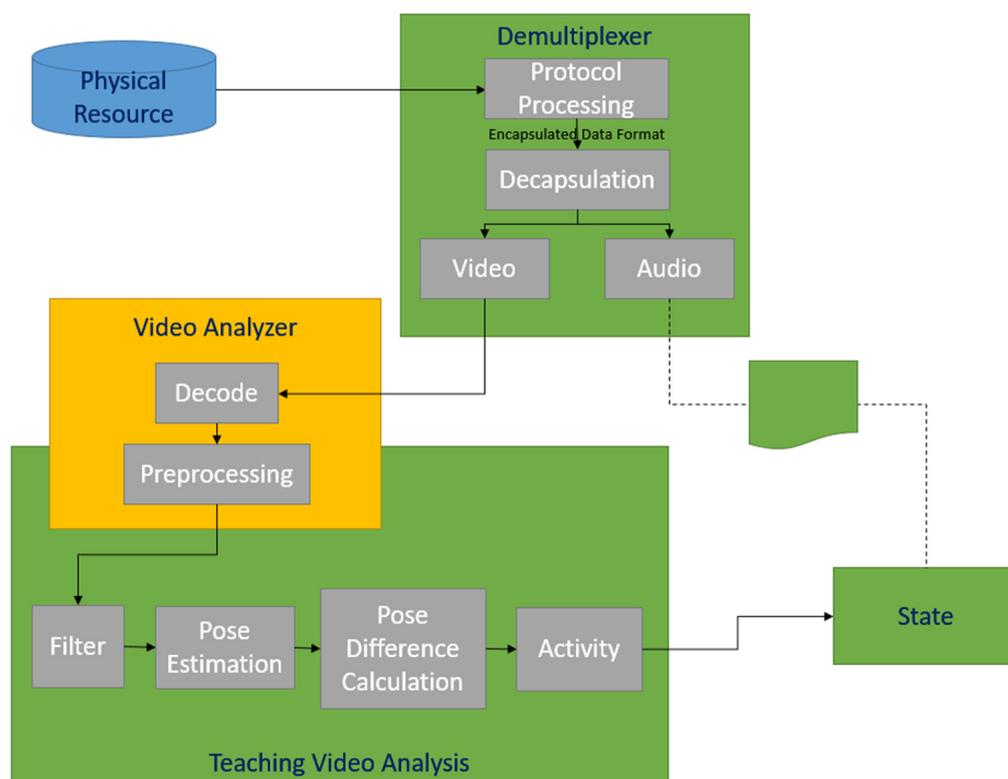


Figure 1. Teaching video analysis flowchart.

After the video analyzer receives the image frame, the analyzed part will select the appropriate filter to process the characters in the classroom environment according to the file settings of the analyzer and then determine the posture of the characters according to the information in the frame of the image. Then, it will iterate to the subsequent image frames to calculate the trainee’s posture difference and finally calculate the trainee’s movement activity during class based on the preset standard.

In the analysis part, pose estimation is a relatively complicated problem. Humans have different body shapes and postures. There are many joints to track (the way joints are expressed in space), and there are often other people or objects around, resulting in visual occlusion.

In the system technology implementation part, there are several methods to choose, such as using special cameras and sensors or using computer vision technology (such as OpenPose or PoseNet) to extract poses from the 2D images to solve the pose estimation problem. Although special cameras and sensors are effective, they often require expensive and not widely distributed technologies and equipment. Therefore, this part of the project was expected to be cheap and easy to obtain by using laptops, webcams, or mobile phone lenses. The trainee’s posture was extracted from the 2D video recorded by the device.

3.2. Recording File Analysis for Speech-to-Text

The main goal of the recording file analysis was to use teaching videos from physical sources and analyze the total amount, proportion, and distribution of speech during class

according to the trainee's audio tracks in the videos and to address further the questioning language (e.g., procedural questioning or inductive question), feedback language, invalid languages such as invalid questions (e.g., duplicates, self-answering, invalid explanations, or digression), and invalid feedback content to judge.

The recording file analysis process is shown in Figure 2. The audio track stream separated in the splitter is input into the audio analyzer for decoding and pre-processing. Hence, the analysis identified what the trainee in the audio stream said. The voice needed to be converted into text before it can be analyzed, and the voice data had continuity. Therefore, the audio stream would be stored as a recording file during pre-processing and then analyzed based on this recording file later.

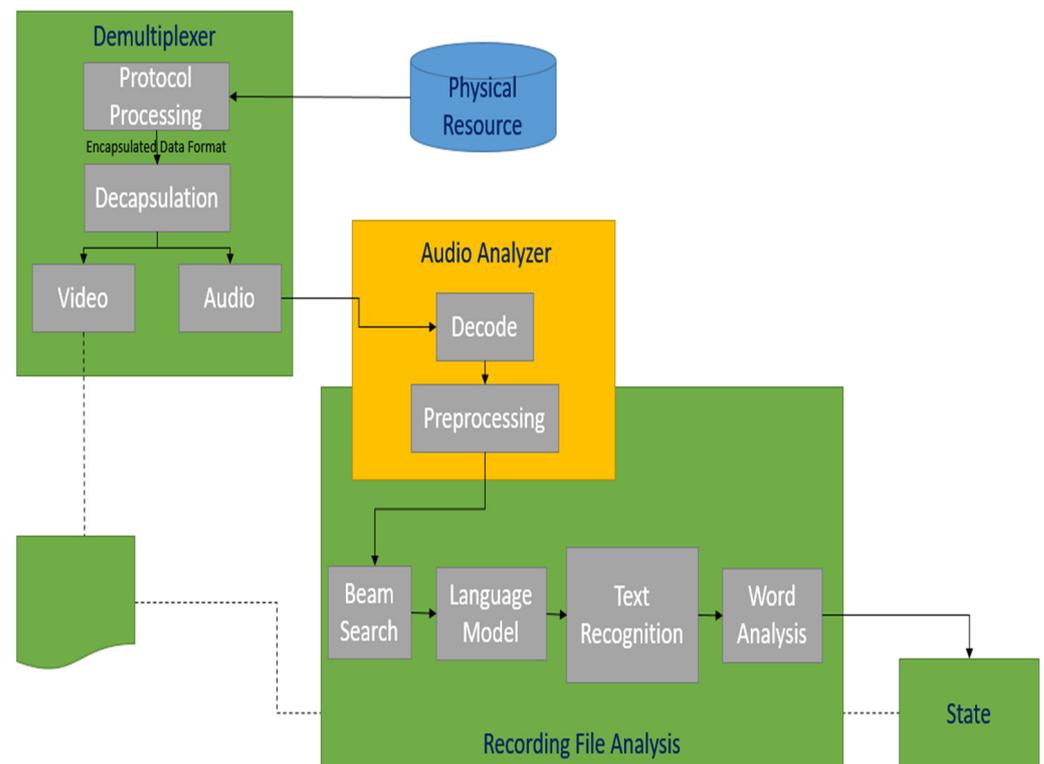


Figure 2. Recording file analysis flowchart.

3.3. Whiteboard Image Analysis for Image-to-Text

The main goal of whiteboard image analysis was to use teaching videos or cloud whiteboard record files from physical sources, analyze the written teaching content of trainees during class according to the content of the whiteboard pictures in the file, and transform them into virtual ones. State judgment included elements such as the relevance of the written content and the curriculum, the recognizability of the written text, and whether the distribution of the written teaching content on the whiteboard was reasonable.

The process of whiteboard image analysis is shown in Figure 3. The source of analysis in this part was more complex than the previous two analyses because the source of the whiteboard data was the physical source of the general whiteboard and included the remote teaching of the course. Virtual sources such as electronic whiteboards or cloud whiteboards may be used. Therefore, the system needed to determine the corresponding whiteboard source during the division of labor. If it came from the whiteboard screen on the teaching video, it would be compared with the whiteboard when the image stream was transmitted with the information and capture the image and pictures on the whiteboard. The whiteboard file would be uniformly converted into a picture format if it came from a virtual electronic whiteboard or cloud whiteboard. The last two sources were the picture format of the whiteboard image and the analysis by the picture. The image analyzer

performed pre-processing in a unified manner and sent it to the subsequent whiteboard image analysis for use.

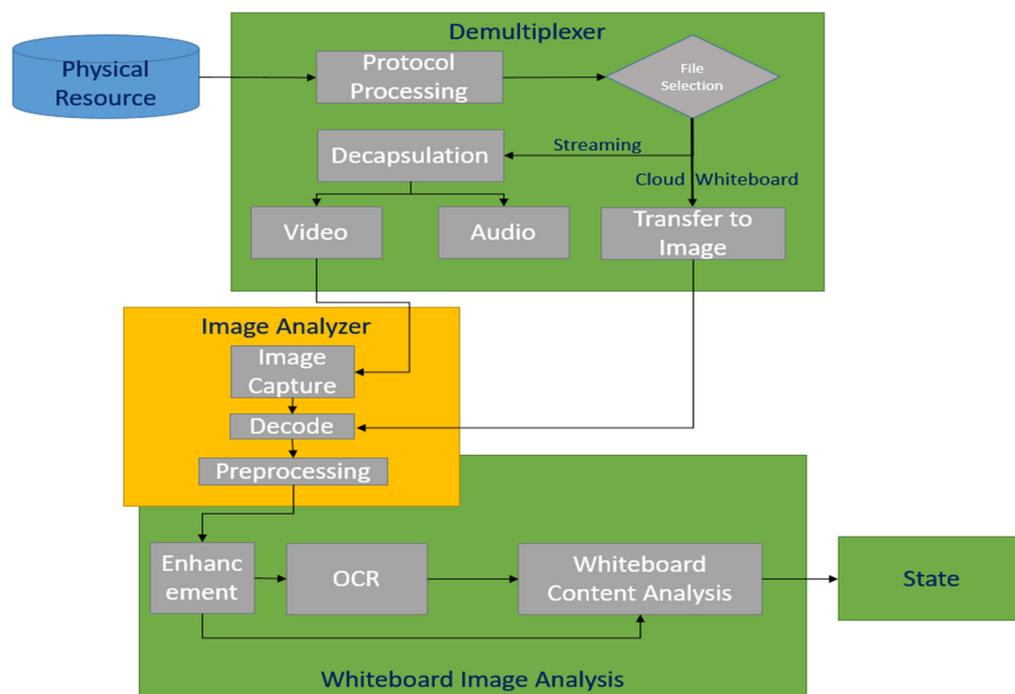


Figure 3. Whiteboard image analysis flowchart.

4. Analog Media to Digital Conversion Method

The purpose of our method was to convert the physical data in the class, such as the video, audio files, and whiteboard images, into objective information. Through this conversion, the advisors can use the information to support their judgment. Our system uses OpenPose [20], the audio, and image recognition to obtain the trainee’s activity and fitness after normalizing. We could analyze the activity by the movement, total amount of speech, the proportion and distribution, and the fitness of procedural questions, introductory questions, failure questions (repetition, self-questioning and self-answering, invalid explanation, or digression), and failure feedback, among other factors. Finally, the activity information could be matched with the current behavior, and the fitness was compared with the established bilingual teaching standards to obtain the feedback of the trainees.

We implemented the method in our online training system, and Figure 4 shows the user interface in our system. The objective information can be observed by the advisor from the right side of the screen. Our method is suited for teacher training courses, but it is more useful for bilingual courses because the advisor not only needs to be concerned about the knowledge in the class but also the correctness of the secondary language. Take Taiwan for example. Our native language is Mandarin, and the secondary language is English. Most of the advisors in the training courses are extremely professional in their approach to the knowledge but not the English. Our system can provide information to help the advisors who are not proficient in this secondary language to give an objective result in a training course.

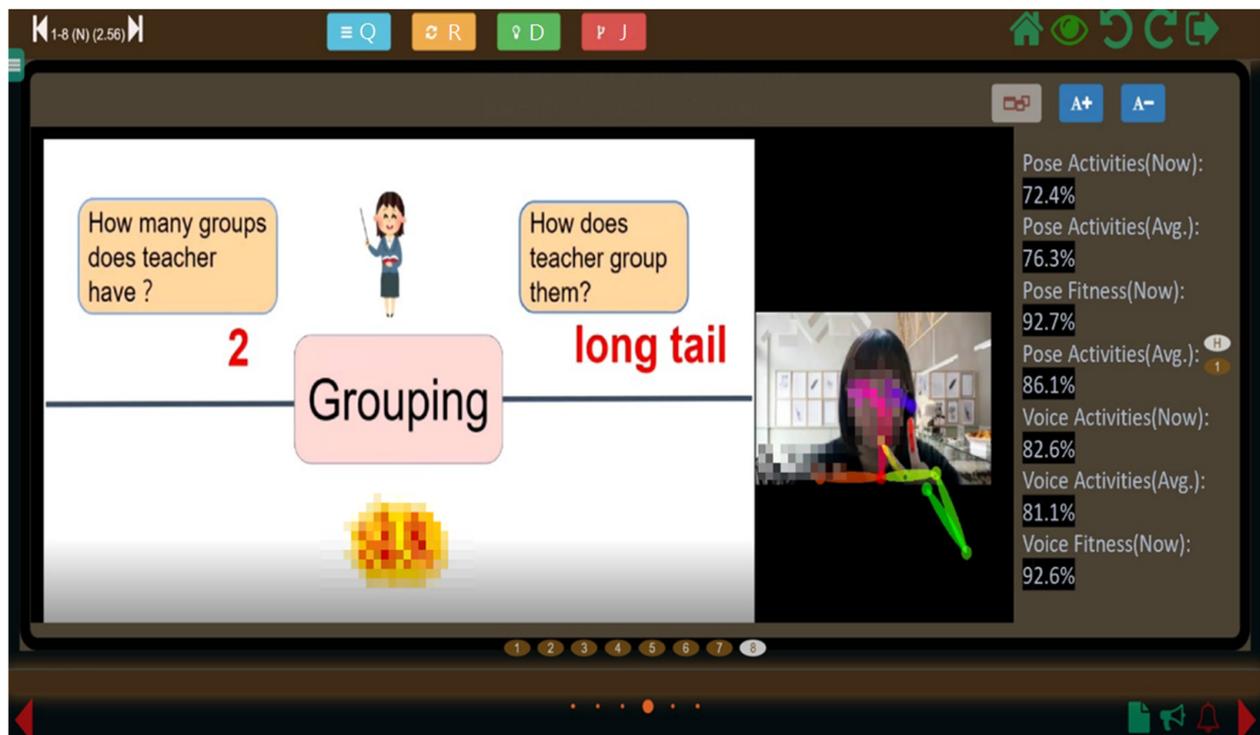


Figure 4. User interface and results for proposed system.

4.1. Pose Difference Calculation

According to the pose estimation result, the first step was generated to change the data from the object to the array when the estimated pose was obtained through the pose estimation technology library, such as PoseNet and OpenPose, PoseNet is built to run on lightweight devices such as a browser or mobile device, whereas OpenPose is more accurate and suitable to be run on a GPU system. The result was a set of data in JSON format, and it represented the position of each key point in the posture. Still, this data representation was slow in analyzing the state, so we expected to compress these values into a one-dimensional array in the analysis as shown in Figure 5. Each of the entries represents the X or Y position of the critical point and some data related to the pose estimation results. As long as we maintained the consistency and predictability of the structure, we could compare the result arrays in almost the same fast way.

Through this part of the analysis, we would be able to transform it into the active state of the trainee through the physical source of the teaching video and pass it to the state attributes in the virtual and real integrated bilingual education training module for subsequent judgments.



Figure 5. Compression of posture estimation data.

4.2. Speech-to-Text Calculation

We used the recognized text to perform subsequent word analysis to determine the critical teaching materials. Teaching materials are the materials used in teaching, also known as teaching content, and they include knowledge, concepts, and all materials used, such as textbooks, exercises, teacher manuals, supplementary materials, test papers, specimens, models, diagrams, audio tapes, video tapes, films, and slides, among others. As shown in Figure 6, the word analysis part received the results from the previous speech converted to text and targeted the recording file analysis process. Various types of analysis were performed on the single words of the content, such as the proportion analysis of the course keywords appearing in the course content, the proportion of the total amount of speech in the course content, the statistics of redundant words, and the calculation of the number of questions. All analyses or judgments would be accelerated by a pipeline. Finally, the analysis results of each part would be counted. Standardized statistical results would be generated according to the audio document settings of the analyzer and sent to the state attribute for follow-up judgment.

By analyzing the recording files in the module, we could finally convert them into the virtual state of the trainees dictating the lecture's content through the physical source of the teaching video and pass this result to the state attribute in the training module, which could help the advisors to make follow-up judgments.

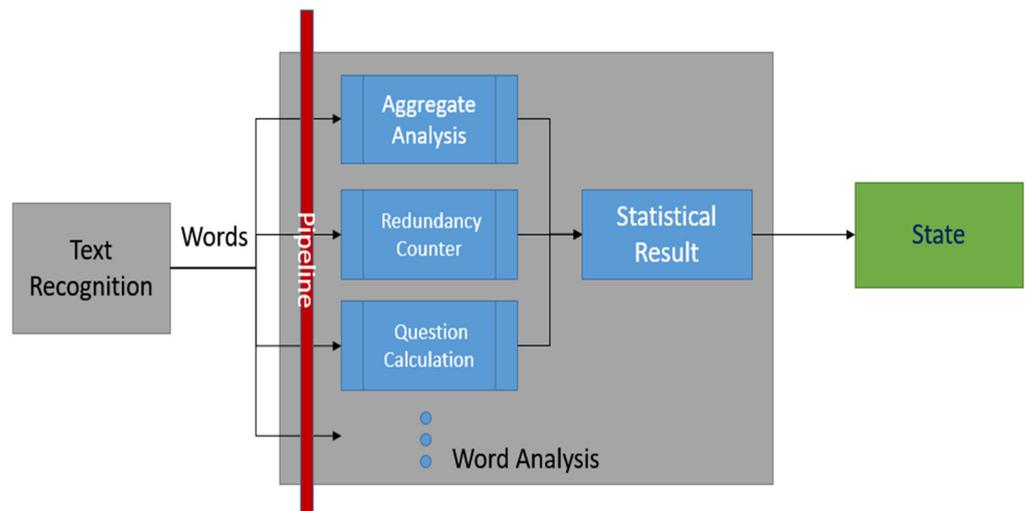


Figure 6. Flow chart of speech conversion text statistics for teaching status results.

Take Figure 7 for example. After the analysis of voice recognition, we would obtain the output result in a text file. Before we calculated the fitness and activity for this training course, we would remove the auxiliary word, the name, and the proper noun in education from the output text file by the pre-defined word table. In this example, we removed “how”, “teacher”, “classification”, and “education” from the output result because they were not necessary for the statistical result in the next step and may have caused mistakes in the calculation of the percentage of aggregate, redundancy, and questions. The final state was the comparison result of the calculation and the original teaching material.

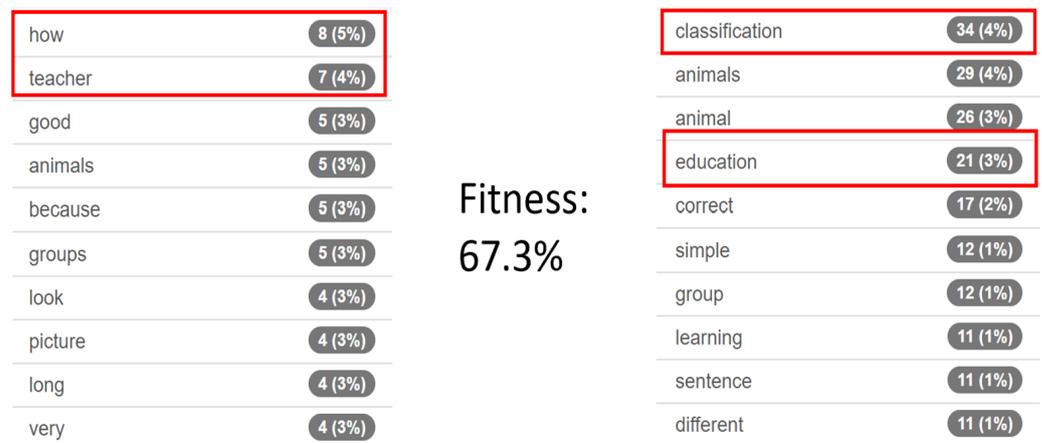


Figure 7. The result of fitness for speech-to-text calculation.

4.3. Image-to-Text Calculation

The whiteboard image analysis process was similar to the recording file analysis process. The difference is that one converts speech to text, and the other converts whiteboard image images into text, as shown in Figure 8. The whiteboard image analysis received the image from after the enhanced whiteboard picture information, and the picture would be processed by optical character recognition to generate the picture and the content of the text. Various analyses would be performed on the text of this content. The whiteboard of the trainee would be obtained directly through the results of the image recognition status and organize the information related to the interactive content of the lecture to generate statistical results and transfer them to the status attribute for subsequent judgment.

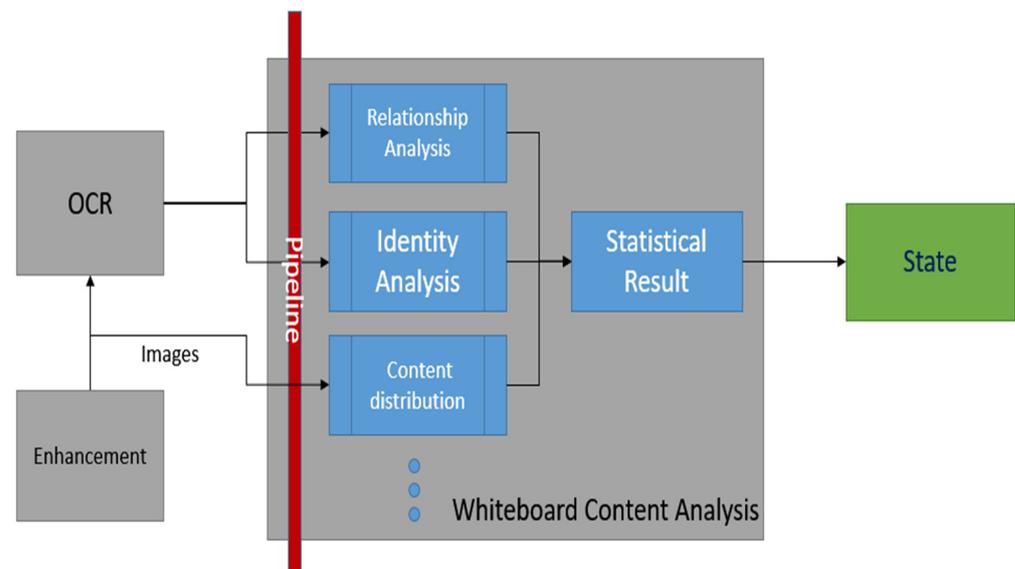


Figure 8. Whiteboard picture conversion text statistics for teaching status result flow chart.

5. Results

5.1. Normalization

Our system is designed to improve online teacher training courses, and bilingual education is a new style for learning in the classroom. It is difficult for our research team to find suitable and pure data of teacher training courses. To find enough cases, we needed to use all the resources we could obtain to prove our system was helpful for the set goal. Our data were from various courses, including math, language, science, society, and art.

The parameters of different courses have a wide range in duration time, key words, content, and number of students, among other factors, and it is hard to judge whether a teacher is “good” and “bad” in a training system. Thus, in this paper, we focused on the objective data from the analysis described in previous sections. Before we evaluated the results, we needed to first normalize the outputs of different courses. For example, in a math course, the teacher may take more time to write the formula on the whiteboard and explain the proof with many sentences, but in an art course, the teacher usually talks less to students and gives them more time to practice, so the drawing data on the whiteboard is also hard to understand.

In the beginning of evaluating the results, the unrecognized output should be abandoned, and the remaining output from various courses will be normalized in a suitable range. For the output of teaching states from the calculations, we performed the Grubbs’s outlier test per course to look for the unrecognized output. We use two-sided Grubbs’s statistics as follows [21]:

$$G > \frac{(K - 1)}{\sqrt{K}} \sqrt{\frac{(t_{\sigma/(2K)}, K - 2)^2}{K - 2 + (t_{\sigma/(2K)}, K - 2)^2}} \tag{1}$$

where G is the test statistic, K is the degrees of freedom of the output data in the analysis, σ is the significance level in the output data of the analysis, and $t_{\sigma/(2K)}, K - 2$ denotes the critical value of the t distribution with $K - 2$ degrees of freedom and a significance level of $\sigma/(2K)$. Figure 9 shows the difference between the original output and the output after the Grubbs’s test as well as the total scores, meaning the sum of the state results calculated from the pose difference, speech-to-text, and image-to-text data.

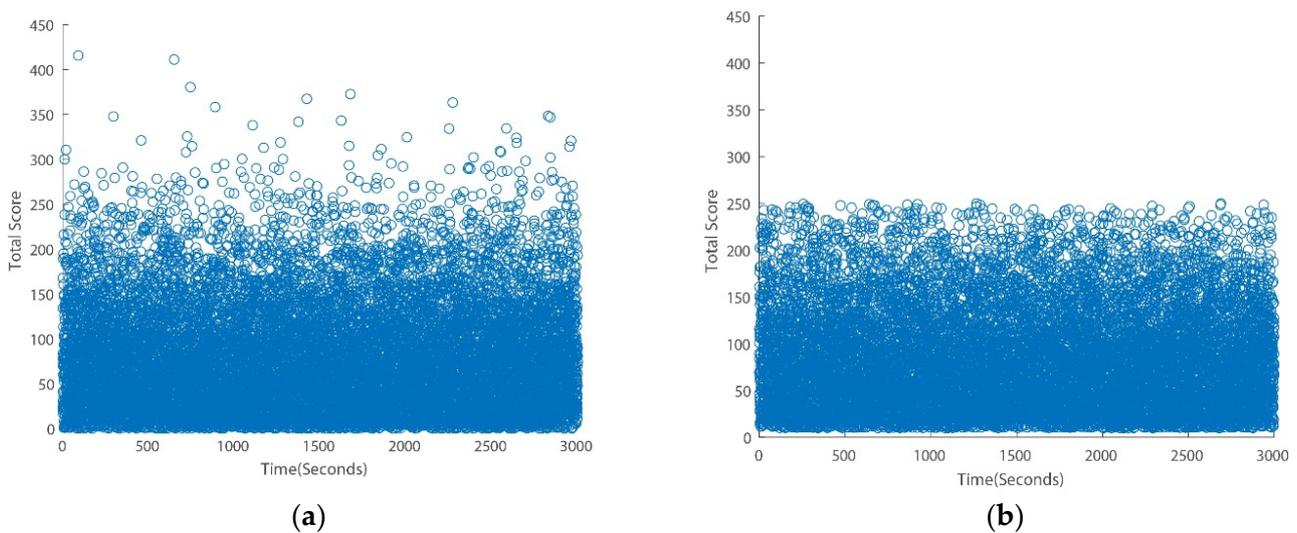


Figure 9. The total calculated score of analysis for a programming course (a) before Grubbs's test and (b) after Grubbs's test.

When we considered the output of the video or audio analysis into consideration, an important thing we needed to be concerned about was each course in various fields having different parameters. For example, a math course may take 1.5 h, and the teacher may use the whiteboard more, but an art course may only take 40 min, and the teacher may usually use body language to describe the concept. Therefore, we also needed to classify the course data into suitable subjects and normalize the time interval and related parameters before we evaluated our system. For normalizing the output raw data to the same interval, and to take the result as the input of the judgment, we used min-max normalization because we used Grubbs's test to restrict the input:

$$O_i = \frac{x_i - \min_{1 \leq j \leq n} \{x_j\}}{\max_{1 \leq j \leq n} \{x_j\} - \min_{1 \leq j \leq n} \{x_j\}} \in [0, 1] \quad (2)$$

where x is the input data from the analysis of the courses and O is the final output for the resulting judgment.

5.2. Evaluation

The data we used to evaluate the results of our method were collected from various teacher training courses for elementary or secondary education. Table 2 shows the detailed information for the courses sampled. We collected the data by recording the streams from classrooms online due to COVID-19. All these courses only had one teacher in the classroom, and both contained a method of immersion, namely content and language integrated learning (CLIL) or English as a medium of instruction (EMI).

Table 2. Sample table of teacher training courses.

Subject	Number of Samples	Input	Duration (Minutes)	Language	Teaching Materials
Math	7	online video	~45	English, Mandarin	textbooks, exercises, and slides
English	6	online video	~30	English, Mandarin	textbooks, exercises, and slides
Science	10	online video	~30	English, Mandarin	textbooks, exercises, and slides
Biology	7	online video	~50	English, Mandarin	textbooks, exercises, and slides
Programming	12	online video	~50	English, Mandarin	exercises and slides
Art	5	online video	~20	English, Mandarin	manuals and models

The first evaluation for our methods was the measurement of the accuracy between different analyses. For the accuracy for the audio and video analysis between the art and

programming courses and the sample results from the manual measurement from the original systems shown in Figure 10, we can observe that the accuracy of the audio analysis was higher than the video analysis. In our opinion, this is because the technique of audio analysis is more mature than that of video analysis. Another phenomenon we can find in this evaluation is the accuracy of the video analysis for art courses being especially low. This issue may be caused by not recognizing the drawn picture on the whiteboard and the unexpected action from the teacher.

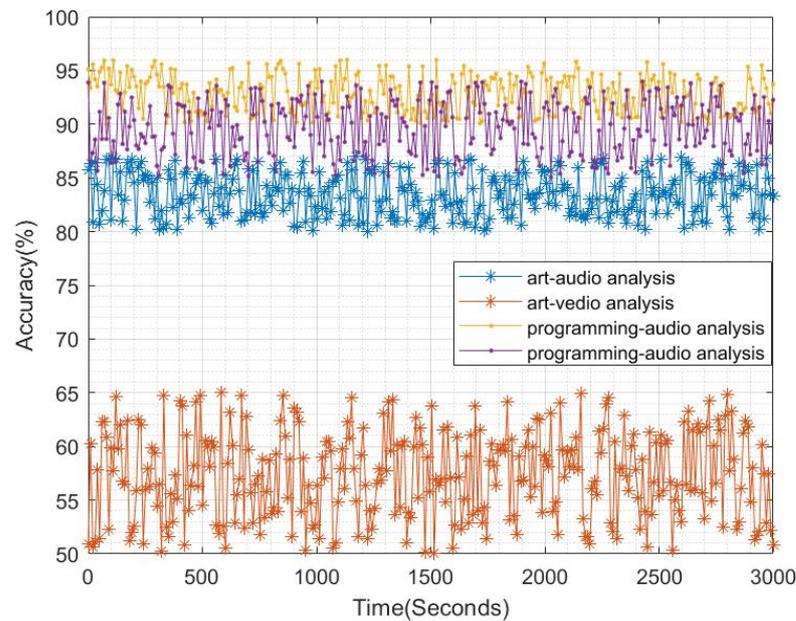


Figure 10. The accuracy for audio and video analysis between the art and programming courses.

The other evaluation concerned the comparison of the total accuracy between different subjects. The sample was also the result of manual measurement from the original courses. From Figure 11, we can find that the total accuracy of the method for most of the teaching subjects was about 70%, meaning that we needed to adjust our method to suit each of the teaching subjects and learning materials. We did this for the programming course, and we can see that the result of programming course was around 97%. Finally, the result of the art course was still undesirable, so we will focus on this issue in further research.

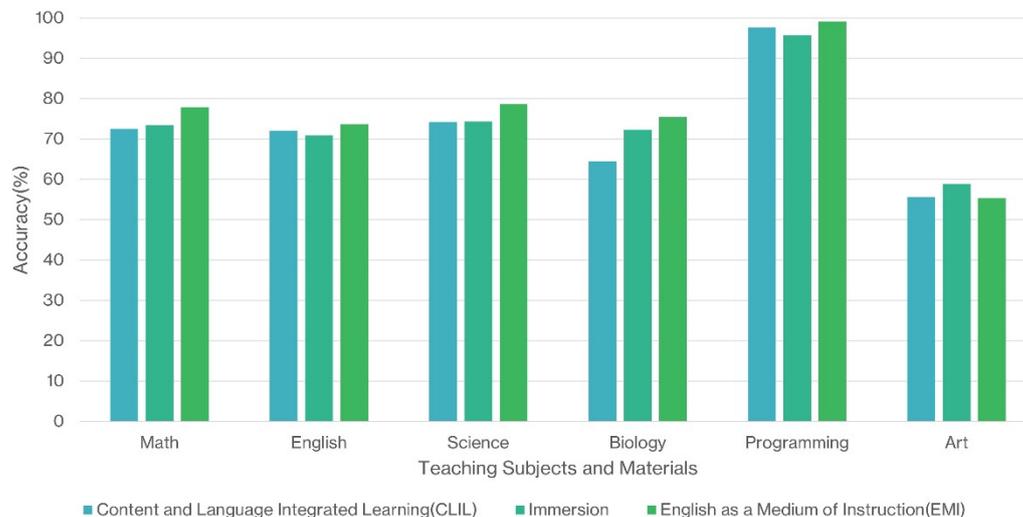


Figure 11. The total accuracy of our method for different teaching subjects and materials.

6. Conclusions

In this paper, we proposed a method for improving the efficiency of online teacher training courses for bilingual education which depends on the technique of CPS. Through pose estimation and video, audio, and image analysis, the proposed methodology can be applied to classroom observation and bilingual teacher certification, assisting in recording and interpreting the teaching demonstration process. The output can facilitate observation of whether or not bilingual teachers can effectively combine subject knowledge with English learning and teaching, including accuracy of language, body movements, textbook preparation, teaching interaction, and English pronunciation and intonation in the classroom. Our evaluation shows that the methods of analysis and calculation had a certain accuracy for most of the teacher training courses compared with manual judgment, and our system can provide objective information for further consideration. We also observed that our model did not perform well in some situations. An example of this was the unrecognized pose estimation in the art courses. We need to record those special poses in the database for further works. Another example was the quick switching between bilingual languages, and we believe that bilingual language identification in quick code-switching speech is a critical issue for future research of bilingual education. This paper can be the basic research for the improvement of online teacher training systems in the future.

Author Contributions: Conceptualization, T.-H.C. and L.-R.C.; data curation, T.-H.C.; formal analysis, Y.-L.C. and J.-S.W.; investigation, Y.-L.C. and J.-S.W.; methodology, T.-H.C. and L.-R.C.; resources, T.-H.C., Y.-L.C., J.-S.W. and R.-G.C.; software, T.-H.C. and L.-R.C.; supervision, T.-H.C.; validation, T.-H.C., C.-Y.S. and R.-G.C.; visualization, T.-H.C., C.-Y.S. and R.-G.C.; writing—original draft, T.-H.C. and C.-Y.S.; writing—review and editing, T.-H.C. and C.-Y.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: We want to thank the budget supported by the National Science Council of Taiwan, R.O.C. under the grant of project MOST-110-2410-H-126-003.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Distance Learning. Myths and Realities. Available online: <https://eric.ed.gov/?id=ED426213> (accessed on 1 January 1998).
2. A Taxonomy of Bandwidth. Myths and Realities. Available online: <http://citeseerx.ist.psu.edu/messages/downloadsexceeded.html#page=76> (accessed on 1 February 1999).
3. Pamela, B.T.; Rosemary, B. Promises and pitfalls of the interactive television approach to teaching adult development and aging. *Educa. Geront.* **1998**, *25*, 741–753.
4. Keskin, S.; Şahin, M.; Uluç, S.; Yurdugul, H. Online learners' interactions and social anxiety: The social anxiety scale for e-learning environments (SASE). *Interact. Learn. Environ.* **2020**. [CrossRef]
5. Robert, M.F.; Chelley, V.; Susan, A.B. E-learning and individual characteristics: The role of computer anxiety and communication apprehension. *J. Comput. Inf. Syst.* **2016**, *46*, 105–115.
6. Guri, R.S. Distance higher education in the digital era: Challenges and prospects. *Dist. Educ. China* **2016**, *6*, 16–25.
7. Paudel, P. Online education: Benefits, challenges and strategies during and after COVID-19 in higher education. *Int. J. Stud. Educ.* **2021**, *3*, 70–85. [CrossRef]
8. Jane, B. Expanding learning opportunities for high school students with distance learning. *Am. J. Distance Educ.* **2014**, *28*, 292–306.
9. Md, A.; Margaret, P. An Innovative Approach toward a Comprehensive Distance Education Framework for a Developing Country. *Am. J. Distance Educ.* **2016**, *30*, 211–224.
10. Olympiou, G.; Zacharia, Z.C. Blending physical and virtual manipulatives: An effort to improve students' conceptual understanding through science laboratory experimentation. *Sci. Educ.* **2011**, *96*, 21–47. [CrossRef]
11. Ian, J.; Malcolm, S.; Alastair, P. Assessing mathematical problem solving using comparative judgement. *Int. J. Sci. Math. Educ.* **2015**, *13*, 151–177.

12. Lei, C.U.; Ka, L.M.; Liang, H.N.; Eng, G.L.; Kaiyu, W. Building an intelligent laboratory environment via a cyber-physical system. *Int. J. Distrib. Sens. Netw.* **2013**, *9*, 109014. [[CrossRef](#)]
13. Miguel, A.R.; Shahriar, M.; Rujia, L. Competitive learning in informatics: The UVa online judge experience. *Olymp. Inform.* **2008**, *2*, 131–148.
14. Szymon, W.; Maciej, A.; Jan, B.; Artur, L.; Tomasz, S. A survey on online judge systems and their applications. *ACM Comput. Surv.* **2018**, *51*, 1–34.
15. An Online Judge System for Beginners. Available online: <https://zerojudge.tw/> (accessed on 21 July 2021).
16. Wang, W.; Han, W.; Chen, W. MetaOJ: A massive distributed online judge system. *Tsinghua Sci. Technol.* **2021**, *26*, 548–557.
17. ZsgsDesign. Available online: <https://github.com/ZsgsDesign/NOJ> (accessed on 4 November 2021).
18. Francisco, H.G.; Jose, D.R.M.; Dovie, A.R.M. The cuban olympiad in informatics: A new stage from the DMOJ online judge. *Olymp. Inform.* **2021**, *15*, 133–141.
19. Qingdao University. Available online: <https://github.com/QingdaoU> (accessed on 18 December 2021).
20. Cao, Z.; Simon, T.; Wei, S.E.; Sheikh, Y. OpenPose: Realtime Multi-Person 2D Pose Estimation Using Part Affinity Fields. *IEEE Trans. Pattern Anal. Mach. Intell.* **2021**, *43*, 172–186. [[CrossRef](#)] [[PubMed](#)]
21. Grubbs, F.E. Procedures for Detecting Outlying Observations in Samples. *Technometrics* **1969**, *11*, 1–21. [[CrossRef](#)]