

KnowLearn: Evaluating cross-subjects interactive learning by deploying knowledge graph

Haolei LIN¹, Junyu CHEN², Hung-Lin CHI^{3*}

¹ Department of Building and Real Estate, Faculty of Construction and Environment, University of The Hong Kong Polytechnic University, Hong Kong, E-mail address: haolei2024.lin@connect.polyu.hk

² Department of Building and Real Estate, Faculty of Construction and Environment, University of The Hong Kong Polytechnic University, Hong Kong, E-mail address: janie-junyu.chen@connect.polyu.hk

³ Department of Building and Real Estate, Faculty of Construction and Environment, University of The Hong Kong Polytechnic University, Hong Kong, E-mail address: hung-lin.chi@polyu.edu.hk

Abstract: In the realm of Architecture, Engineering, and Construction (AEC) education, various factors play a crucial role in shaping students' acceptance of the learning environments facilitated by visualization technologies, such as virtual reality (VR). Works on leveraging the heterogeneous educational information (i.e., pedagogical data, student performance data, and student survey data) to identify essential factors influencing students' learning experience and performance in virtual environments are still insufficient. This research proposed KnowLearn, an interactive learning assistant system, to integrate an educational knowledge graph (KG) and a locally deployed large language model (LLM) to generate real-time personalized learning recommendations. As the knowledge base of KnowLearn, the educational KG accommodated multi-faceted educational information from twelve perspectives, such as the teaching content, students' academic performance, and their perceived confidence in a specific course from the AEC discipline. A heterogeneous graph attention network (HAN) was utilized to infer the latent information in the KG and, thus, identified the perceived confidence, intention to use, and performance in a relevant quiz as the top three indicators that significantly influenced students' learning outcomes. Based on the information preserved in the KG and learned from the HAN model, the LLM enhanced the personalization of recommendations concerning adopting virtual learning environments while protecting students' privacy. The proposed KnowLearn system is expected to feasibly provide enhanced recommendations on the teaching module design for educators from the AEC domain.

Key words: Learning Assistant System, Knowledge Graph, Large Language Model, Visualization Technologies

1 INTRODUCTION

1.1 Visualization Technologies in Architecture, Engineering, and Construction

University education plays a vital role in modeling future industry-related talents in the realm of Architecture, Engineering, and Construction (AEC). In recent years, innovative visualization technologies, such as virtual reality (VR), have been widely used in the AEC industry [1,2]. It is crucial to improve the acceptance of visualization technologies and the efficiency of learning/training procedures among students for an effective learning experience. VR has been utilized to enhance students' understanding of domain knowledge, such as demonstrating spatial arrangement of structures, hazard identification at construction sites, spatiotemporal modeling of construction processes, and so forth [3,4,5]. However, this does not mean that all students can benefit equally from learning by using visualization technologies.

Cave automated virtual environment (CAVE), as one of the representative VR implementations for instance, is an innovative system that can be utilized to provide immersive learning environments [1,6]. As a highly immersive VR system, CAVE can facilitate face-to-face or mixed communication between teachers and students, which is suitable for small-scale group learning. However, together with other innovative visualization techniques, one primary challenge in garnering effective learning is properly refining the teaching modules per the developed virtual model. Another difficulty is related to the impacts of psychological factors, such as students' cognitive load and technology acceptance, which should be carefully considered during the implementation of such teaching innovations. These issues hinder the widespread adoption of innovative visualization technologies in AEC educational teaching and learning activities. Previous research has shown that factors such as perceived facilitating conditions and perceived effort expectancy would affect students' acceptance of VR [7]. This study uses the knowledge graph (KG) technique, an effective knowledge management tool, to combine domain knowledge with semantic information to facilitate the holistic understanding of AEC education features. The KG could extract information, map knowledge content, and obtain the main factors influencing students' VR use.

1.2 Applications of Large Language Models

ChatGPT, a variant of the generative pre-trained transformer 3 (GPT-3 model), is a large pre-trained language model launched by OpenAI in November 2022 and opens a new booming era of generative AI study. It is a deep learning-based natural language processing (NLP) chatbot based on the Transformer architecture and leverages both supervised and unsupervised learning to obtain language knowledge from massive text data [8]. ChatGPT has robust language understanding and generative capabilities to generate coherent, grammatically sound, and fluent responses to a range of prompts and queries [9,10]. ChatGPT and other large language model (LLM)-based chatbots reached average performance or outperformed in a series of tests, such as MBA, law exams, and medical exams [9].

LLM's powerful understanding and generative capabilities have also been widely used in various fields. Many developers have developed LLM-based chatbots that are specifically suitable for their respective professional fields, providing more efficient and targeted help to people in different industries from multiple perspectives. In daily life, by connecting to external websites, these agents (chatbots) can better understand the intentions of network users and help them search and find the information they want to obtain more efficiently [11]. Many software developers use LLM-based chatbots at work to help complete computers and check related code issues. Many literary and artistic workers can also use agents to assist in their creative work. Students can use them to organize and understand course content in their studies; researchers can also get help from LLM-based chatbots in document management and content search. Based on LLM-based chatbots, this study uses their powerful understanding and generative capabilities to generate learning suggestions for students.

2 OBJECTIVES

The study aims to provide a system called KnowLearn to construct and utilize the Knowledge Graph (KG) to enhance the student learning experience in virtual learning environments. The proposed KnowLearn system endeavors to tackle the primary lack-of-support issues arising from applying innovative visualization tools (e.g., the CAVE) and large language models (e.g., ChatGPT) in tertiary AEC education. Based on the above background, this paper proposes the following objectives:

- 1) To construct a framework of an educational KG for interactive learning in the AEC discipline based on the pedagogical data, student feedback, and performance.

- 2) To develop the Knowlearn platform where the pre-computed KG is visualized, and AEC students can receive personalized learning recommendations through a locally deployed LLM.
- 3) To testify the proposed platform through an undergraduate course in construction, further evaluating the recommendation results for learning.

3 METHODOLOGY

Initial data were collected through questionnaires from undergraduate subjects at the AEC discipline. A KG was established after data preprocessing to summarize the relevant factors that affected student performance. Then, the weights of factors affecting student performance were obtained to compare and analyze the results. In addition, a corresponding LLM was deployed locally to provide personalized suggestions. It ensures that important and private information about students and teaching activities will not be leaked. The KnowLearn was built by combining the weights of different factors obtained above and the LLM to evaluate student performance.

3.1 Data Collection

The data was collected from a university subject regarding introducing construction technology and materials. The objectives of this course are to provide students with an understanding of 1) the functional requirements of buildings; 2) how different building elements and components behave; 3) to realize the range of building materials available for construction; 4) to gain an understanding of the key design concepts in buildings; and 5) to support students with better understanding on the construction technology and material through CAVE interactive learning.

The data obtained in this research came from a total of 107 undergraduate students who chose this subject. Students have experienced a series of training in operating VR devices. Students in groups received guidance in the CAVE environment, which provided a virtual construction scene to help them interact with building elements, further recognizing their functions and properties. Data were collected in the form of questionnaires afterward, which aimed to acquire feedback about the CAVE-enabled learning experience and mainly included four aspects: Basic Information, IT Application, Factors the respondent thinks Influence Learning, and Comments. It is worth noting that IT Applications include three aspects: Perceived Ease of Use, Perceived Usefulness, and Intention to Use. Factors that Influence Learning include a total of 10 factors, such as fidelity and presence. **Figure 1** shows the components of the questionnaire.

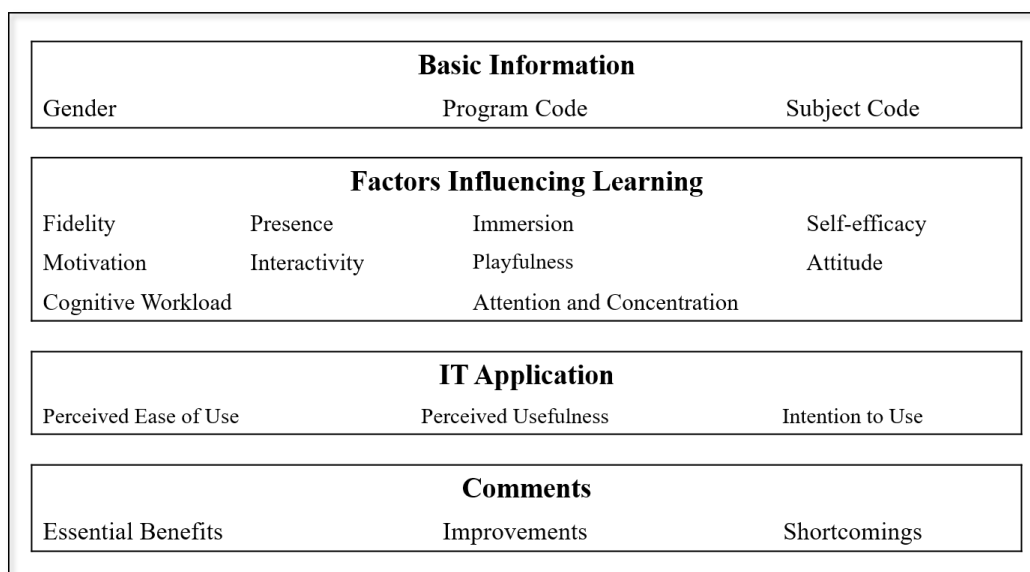


Figure 1. Questionnaire Design

3.2 Heterogeneous Graph Attention Network

A heterogeneous graph consists of a set of nodes V and a set of edges ε . There is more than one type of node and edge, so the sum of categories is greater than 2. A meta-path is a compound relationship connecting two objects and is a widely used semantic capture structure. Depending on the meta-path, the relationships between nodes in heterogeneous graphs can have different semantics. The heterogeneous graph attention network (HAN) is a heterogeneous graph neural network based on the attention mechanism. It can extract richer semantic information from complex graph structures [12]. The HAN model can simultaneously extract node-level attention and semantic-level attention. Node-level attention aims to learn the importance between a node and its meta-path-based neighbors, while semantic-level attention can learn the importance of different meta-paths.

By analyzing and preprocessing the collected data, 80 valid questionnaires were obtained to create the meta-path of 'Student-Feature-Student' based on students' feedback. A total of 50 features were used to measure factors, such as Perceived Ease of Use, and construct a HAN model.

3.3 Locally Deployed Large Language Model

In order to improve the capabilities of LLMs in professional fields, this research collected and integrated pedagogical principles, learning suggestions, and other knowledge related to the AEC discipline. These textual data were used to assist locally deployed LLMs with a specific understanding and generative capabilities of AEC disciplines. Locally deployed LLM helps provide Q&A interactions while protecting the privacy and security of students' data, achieving instant personalized learning recommendations acquisition.

3.4 KnowLearn System

KnowLearn integrates the KG constructed by the HAN model and the locally deployed LLM. This system utilizes heterogeneous data to extract significant factors influencing students' learning performance. The generative capabilities of LLM are utilized in the system by providing personalized learning recommendations and potentially promoting the improvement of teaching module design in the future. **Figure 2** shows the user interface (UI) of the KnowLearn system. The content in the blue box is an introduction to the KnowLearn system and a description of the data collection format. The content in the red box collects user preferences and links the knowledge graph trained by the HAN model. The content in the green box is the individual learning suggestions generated by LLM that are customized to the students.

When users interact with this system, they can start by reading the contents of the introduction. It provides brief information on the KnowLearn system and explains how to answer the questions in the following section (red box area), on which a scale of 1 to 7 is used, where 1 represents "strongly disagree," and 7 represents "strongly agree." Afterward, users provide responses by entering the corresponding feedback under each question. Once all the required questions have been finished, click the 'Submit' button at the bottom, and there will be personalized recommendations to be generated following below.

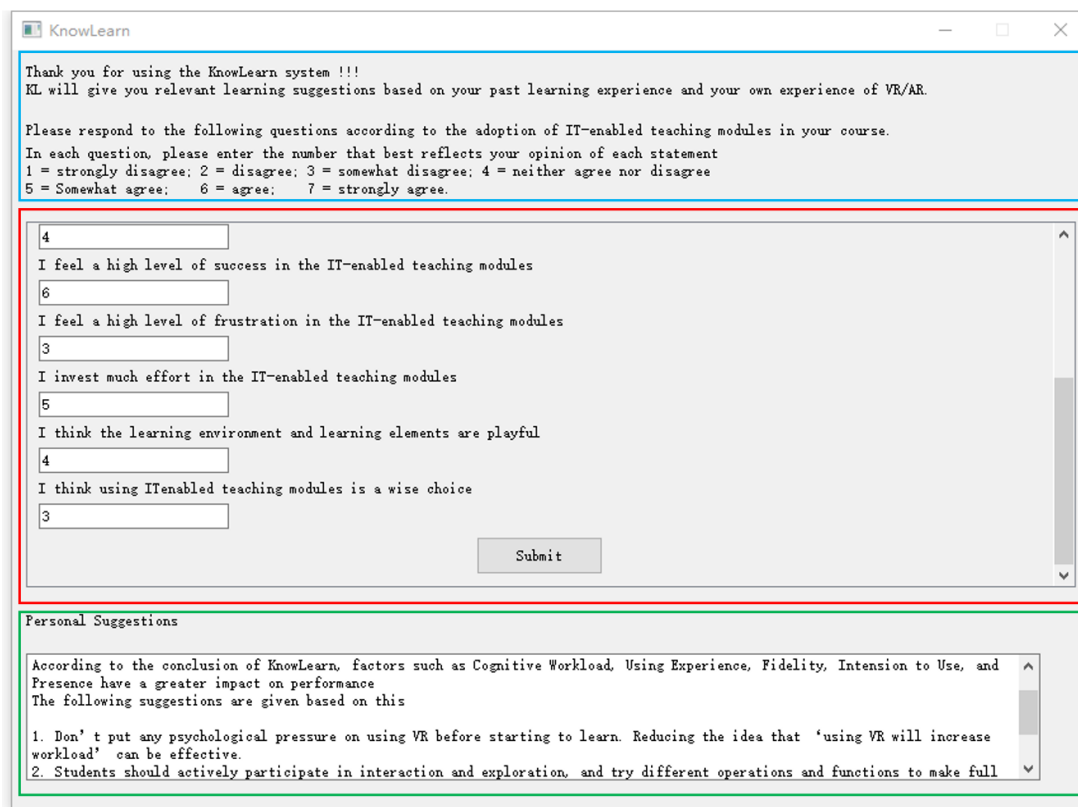


Figure 2. The UI of the KnowLearn System

4 RESULTS

The HAN model was used to extract significant factors affecting students' learning performance. The model's accuracy through validation reached approximately 93.3%, and **Figure 3** shows the detailed information. It was found that the factors in **Table 1** significantly impacted students' performance. Although there were a total of 50 different questions in the questionnaire, the 15 most representative ones accounted for approximately 95.3% of the total. The factor that had the greatest impact on student performance was 'Cognitive Workload,' which accounted for 44.28%. 'Using Experience,' 'Fidelity,' 'Intention to Use,' and 'Presence' accounted for 9.14%, 8.37%, 8.10%, and 7.01% respectively. The factors that had relatively little impact on student performance were 'Perceived Ease to Use,' 'Immersion,' 'Attention and Concentration,' 'Interactivity,' and 'Others,' accounting for no more than 1%.

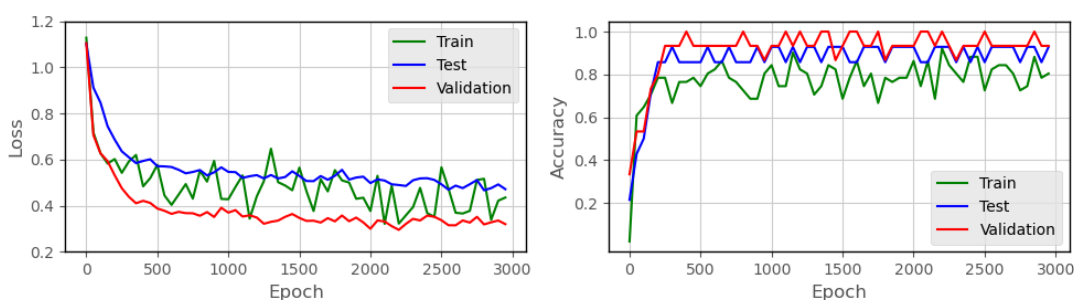


Figure 3. Results of the HAN Model

The factors corresponding to questionnaire questions were extracted by the HAN model and collected again in the KnowLearn system to help users get personalized learning recommendations in real-time. This study simulated two prompts that students may raise and compared the responses from the ChatGPT and KnowLearn systems, respectively. Firstly, users directly asked for some learning

suggestions. **Table 2** displays ChatGPT's prompt, corresponding responses, and KnowLearn's personalized responses. There was a larger amount of responses generated by ChatGPT, but the suggestions were less specific and informative. The fine-tuned KnowLearn system retained specific information about the subject and gave specific personalized answers based on the objectives of the subject, the composition of grades, and other characteristics.

Table 1. Weight of Factors

| Factors | Weights | Factors | Weights |
|--------------------|---------|-----------------------------|---------|
| Cognitive Workload | 44.28% | Motivate | 4.82% |
| Using Experience | 9.14% | Self-Efficacy | 2.49% |
| Fidelity | 8.37% | Perceived Usefulness | 2.10% |
| Intention to Use | 8.10% | Perceived Ease to Use | 0.72% |
| Presence | 7.01% | Immersion | 0.44% |
| Playfulness | 5.71% | Attention and Concentration | 0.37% |
| Attitude | 5.43% | Interactivity | 0.36% |

Table 2. Prompt and Responses from ChatGPT and KnowLearn: Subject Suggestions

| | |
|-----------|---|
| Prompt | I will study a subject called 'Introductory Construction Technology and Materials.' Please give me some suggestions. |
| ChatGPT | <p>Certainly! Here are some suggestions to help you with your study of "Introductory Construction Technology and Materials":</p> <ol style="list-style-type: none"> 1. Familiarize yourself with the course syllabus: Take the time to thoroughly review the course syllabus, as it will provide you with a clear understanding of the topics covered, learning objectives, and assessment methods. This will help you plan your study schedule and stay organized throughout the course. 2. Actively engage with the course materials: Make sure to actively engage with the assigned readings, lecture notes, and any supplementary materials provided by the instructor. Take detailed notes, highlight key concepts, and ask questions to deepen your understanding of construction technology and materials. <p>...</p> <ol style="list-style-type: none"> 8. Stay updated with industry advancements: Construction technology and materials are constantly evolving. Stay updated with the latest industry trends, advancements, and new materials. Follow industry publications, attend seminars or webinars, and join relevant professional organizations to stay informed about the latest developments in the field. <p>Remember to maintain a consistent study schedule, stay organized, and seek help whenever needed. With dedication and active engagement, you can excel in your study of "Introductory Construction Technology and Materials."</p> |
| KnowLearn | <p>According to KnowLearn, here are some suggestions:</p> <ol style="list-style-type: none"> 1. The objectives of the course are to understand the functions of buildings, how different building elements and components interact, the classification and properties of building materials, etc. You can focus on learning based on objectives. 2. The written examination and tutorial assessment scores account for 60% and 15% respectively. Searching past test papers from the school library's website can help review better. 3. This subject will have Laboratory Sessions to help students better understand the learning content. Active participation can help you better understand and retain what you have learned. |

Secondly, **Table 3** displays ChatGPT's prompt and corresponding responses when students had concerns about the use of VR technology and the comparison of KnowLearn's personalized suggestions, respectively. ChatGPT's responses were still general, while KnowLearn's personalized suggestions, which were based on the information extracted from the knowledge graph, were more professional and specific from different perspectives and reflected essential factors, such as 'Cognitive Workload,' 'Playfulness' and so forth. Thought a solid solution for the student to ease pressures further would need further data support by improving the KG other than factors. These would be the future direction of the following research works.

Table 3. Prompt and Responses from ChatGPT and KnowLearn: VR Concern

| | |
|-----------|--|
| Prompt | I have some concerns about using VR to learn about a subject called 'Introductory Construction Technology and Materials'. Please give me some suggestions for using VR. |
| ChatGPT | <p>Using VR for learning about introductory construction technology and materials can be a highly effective and immersive way to enhance your understanding of the subject. Here are some suggestions for utilizing VR in your learning process:</p> <ol style="list-style-type: none"> 1. Virtual Site Tours: VR can simulate real construction sites, allowing you to explore different environments and observe construction processes. You can virtually visit construction sites, examine materials, and understand how different components and technologies are used in practice. 2. Interactive Demonstrations: VR can provide interactive demonstrations of construction techniques and material applications. You can virtually engage with tools, equipment, and materials, allowing for hands-on learning experiences without the physical limitations. ... 10. Real-time Updates: VR platforms can be updated with the latest construction technologies and materials, ensuring that you stay up to date with industry advancements. <p>Remember to complement your VR learning with traditional study materials, textbooks, and practical experiences whenever possible. VR should be used as a supplementary tool to enhance your learning journey.</p> |
| KnowLearn | <p>According to KnowLearn's judgment, you have certain concerns about using VR for learning. According to the conclusion of KnowLearn, factors such as 'Cognitive Workload,' 'Using Experience,' 'Fidelity,' 'Intension to Use,' and 'Presence' have a greater impact on performance. The following suggestions are given based on this:</p> <ol style="list-style-type: none"> 1. Don't put any psychological pressure on using VR before starting to learn. Reducing the idea that 'using VR will increase workload' can be effective. (Cognitive Workload) 2. You can try different operations and functions to make full use of the learning opportunities, you can try to operate and assemble architectural elements by yourself to increase the fun of learning. (Playfulness) 3. If you have never used VR in the past or rarely used related equipment, don't put any special psychological pressure on it. Usually, the requirements for students to use VR and other equipment in teaching are very low, and it is very easy to learn. (Using Experience) |

After fine-tuning the LLM with the relevant dataset of AEC education, it was observed that the suggestions provided became more specific with the identified influential factors based on the student feedback on this subject. The KnowLearn system offered more relevant recommendations based on the factors with greater weight determined by the HAN model, while particular "actions" the student could take and collective improvement direction for the subject design need further exploration.

5 CONCLUSION

With the development of advanced visualization technologies, such as virtual reality (VR), it is essential to consider the adoption efficiency and effectiveness in the university education of the AEC area. While visualization technologies improve teaching immersion and interactive experience, they also increase some students' learning concerns. In this study, a personalized learning recommendation system called KnowLearn, was developed. It consists of a domain-specific knowledge graph (KG) generated by a heterogeneous graph attention network (HAN) model, helping to identify the impacts of the influential factors on AEC students' learning performance through VR technology adoption. Furthermore, the system provides a Q&A interactive function to offer personalized learning recommendations through a locally deployed large language model (LLM). Further analysis of VR education applications based on the subject of construction technology and material was conducted on undergraduate students in AEC fields. The analysis conducted in this study identified important factors. 'Cognitive Workload' is the most important factor to influence students' performance, with a single accounting for 44.28%, indicating that the amount of learning content from the student's perspective can significantly affect performance. Compared with traditional education, more immersive devices such as VR may help students concentrate while increasing their mental and physical pressure. The results

indicated that although a high concentration of attention can help one study more efficiently, the learning time cannot be sustained due to the accumulation of mental and physical fatigue. Therefore, in the design of relevant courses, it may be necessary to reduce the teaching time compared with traditional teaching, but what will be gained is students' more profound understanding and impression of pertinent knowledge.

Factors such as 'Using Experience,' 'Fidelity,' 'Intention to Use,' 'Presence,' 'Playfulness,' and 'Attitude' also affect students' performance. The weight of each factor is above 5%. 'Fidelity,' 'Presence,' and 'Playfulness' are all related to the immersion of VR, which are the important features that distinguish VR-enabled teaching from traditional teaching. 'Intention to Use' and 'Attitude' reflect students' attitudes towards learning to some extent, showing that stimulating students' interest is the most important way to help students improve their performance.

The significance of this study is that the KnowLearn system and corresponding analysis proved their feasibility in determining the weighted factors that affect student performance at the AEC undergraduate education. The system can further provide personalized recommendations through the combination of LLM's generative capabilities and analytical results. It also provides a potential idea for the design of subsequent AEC subjects and helps VR and other visualization technologies to be better applied in higher education. Further studies, including improvement and validations of the usefulness of personalized recommendations, should be conducted to justify the usage and reliability of the proposed system.

ACKNOWLEDGMENTS

The authors would like to thank The Hong Kong Polytechnic University for the funding support under the Teaching Development Grant (No. TDG22-25/SEI-1).

REFERENCE

- [1] Li X, Yi W, Chi H L, et al. A critical review of virtual and augmented reality (VR/AR) applications in construction safety[J]. *Automation in Construction*, 2018, 86: 150-162.
- [2] Wang, P., Wu, P., Wang, J., Chi, H. L., & Wang, X. (2018). A critical review of the use of virtual reality in construction engineering education and training. *International journal of environmental research and public health*, 15(6), 1204.
- [3] Fogarty, J., McCormick, J., & El-Tawil, S. (2018). Improving student understanding of complex spatial arrangements with virtual reality. *Journal of Professional Issues in Engineering Education and Practice*, 144(2), 04017013.
- [4] Perlman, A., Sacks, R., & Barak, R. (2014). Hazard recognition and risk perception in construction. *Safety Science*, 64, 22-31.
- [5] Wong, K. D. A., Wong, K. W. F., & Nadeem, A. (2011). Building information modeling for tertiary construction education in Hong Kong. *Journal of Information Technology in Construction (ITcon)*, 16(27), 467-476.
- [6] Zheng, L., Zhu, M., & Yu, H. (2020, August). A virtual environment making method for CAVE system. In *2020 International Conference on Computing and Data Science (CDS)* (pp. 377-380). IEEE.
- [7] AL-Oudat, M., & Altamimi, A. (2022). Factors influencing behavior intentions to use virtual reality in education. *International Journal of Data and Network Science*, 6(3), 733-742.
- [8] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., ... & Polosukhin, I. (2017). Attention is all you need. *Advances in neural information processing systems*, 30.
- [9] Ray, P. P. (2023). ChatGPT: A comprehensive review on background, applications, key challenges, bias, ethics, limitations and future scope. *Internet of Things and Cyber-Physical Systems*.
- [10] Lo, C. K. (2023). What is the impact of ChatGPT on education? A rapid review of the literature. *Education Sciences*, 13(4), 410.
- [11] Bing Chat. <https://www.digitaltrends.com/computing/how-to-use-microsoft-chat-gpt-bing-edge/>, 2023. Available Online, Accessed on March.
- [12] Wang, X., Ji, H., Shi, C., Wang, B., Ye, Y., Cui, P., & Yu, P. S. (2019, May). Heterogeneous graph attention network. In *The world wide web conference* (pp. 2022-2032).