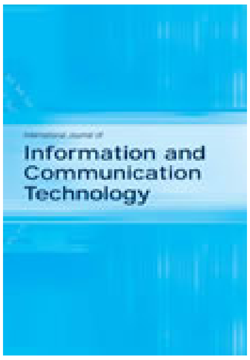


International Journal of Information and Communication Technology

ISSN online: 1741-8070 - ISSN print: 1466-6642
https://www.inderscience.com/ijict

Real-time AI-regulated animation-user interaction system in virtual reality environments

Jun Liu

DOI: 10.1504/IJICT.2026.10076062

Article History:

Received:	23 September 2025
Last revised:	20 November 2025
Accepted:	02 December 2025
Published online:	16 February 2026

Real-time AI-regulated animation-user interaction system in virtual reality environments

Jun Liu

College of Art,
Cheongju University,
Cheongju, 28503, South Korea
Email: junliu9409@163.com

Abstract: This study presents a real-time AI-regulated animation and user interaction system that leverages machine learning to enhance immersion in virtual reality (VR) environments. The framework integrates real-time simulation, CAD optimisation, and AI-driven animation to deliver responsive, realistic, and user-friendly interactions. Although challenges remain in resource utilisation and frame rate stability, experimental evaluations demonstrate high accuracy, responsiveness, and usability. The findings suggest that AI-governed VR systems hold significant potential in education, healthcare, and training for high-risk environments. As VR adoption expands across medicine, education, and the arts, the need for machines capable of dynamically controlling interactions and animations becomes critical for achieving presence and adaptability. Earlier rule-based and tele-operated approaches lacked realism and scalability, while newer AI-powered methods offer greater flexibility yet still face integration challenges. This system addresses these limitations by combining AI models with CAD-optimised animations, ensuring speed, precision, and usability in real-time.

Keywords: augmented reality; AR; data mining algorithms; interaction with users; AI-regulated rendering; computer-aided design; CAD; optimisation; immersive reality; VR.

Reference to this paper should be made as follows: Liu, J. (2026) 'Real-time AI-regulated animation-user interaction system in virtual reality environments', *Int. J. Information and Communication Technology*, Vol. 27, No. 11, pp.38–61.

Biographical notes: Jun Liu is currently a PhD candidate at Cheongju University, South Korea. His main research interests lie in the interdisciplinary integration of artificial intelligence, art, and animation, with a particular focus on human-computer interaction and AI-driven digital art and animation creation technologies.

1 Introduction

Reasonably priced head-mounted displays (HMDs) that can provide extremely strong environments for creating applications, such as unreal engine and unity, and engaging in immersive experiences have propelled virtual reality (VR) to the forefront of mainstream interaction technologies in recent years (Pellas et al., 2020). The tourism, education, and health and rehabilitation industries are just a few that have benefited from VR experiences. Despite variations in intended audience, content, interaction style, and hardware, all VR experiences aim to make users feel physically present. A person is said to be in a 'mental condition or individual viewpoint' when they fail to properly recognise the impact of technology on their lived experience, even though some of what they are experiencing right now is shaped by or generated by technological developments that humans have created. In other words, the user loses track of the fact that they are participating in a technologically-mediated simulation and becomes fully immersed in the virtual world. As a result, one of the biggest obstacles in creating VR experiences is making presence possible (Lorentz et al., 2021). Information quality, user agency, realism, and specific psychological, emotional, and behavioural aspects are among the many components that have been shown in prior studies to contribute to presence. Possible factors that evoke a feeling of immersion in VR include top-notch gear like VR headsets, devices for controlling and tracking, devices for activities, and accurate code that orchestrates the experience. We call these things 'presence factors', and we mean everything about a VR system or its use environment that helps to evoke a sense of being there (Laine and Suk, 2024).

Research on gesture-based interactions began in the early days of VR to make the technology more user-friendly and intuitive. A gesture-based interface allows the user to control the computer by moving their hands or other limbs. To achieve interface transparency in HMD-based VR, one widespread implementation of the larger idea of a natural user experience (NUI) is a gesture-based interface. There has been tremendous expansion in the VR and augmented reality industries due to the rising consumer desire for immersive technologies. These markets are expected to grow by an additional $16.1 billion in the near future, surpassing the original projection of $1.6 billion by 2025. In 2020, they reached $18.8 billion. New philosophical debates on how human function in scientifically mediated situations have emerged in response to the fast growth in VR technology, which has altered human interaction with one another and the environment (Abramczuk et al., 2023). VR has moved beyond its initial application in games and entertainment to become a potent tool for recreating real-world experiences, opening the door to new ways of interacting with one another that are not limited by distance or time (Archer et al., 2022). As a result of this transformation, people can have conversations that seem remarkably similar to those they would have in real life, which is a huge step forward in the evolution of human connection.

Virtual interactions are getting more complex all the time due to technological advancements. Initially, the internet just allowed text-based communication, but as time went on, it adapted to enable video and audio conversations as well (Reynolds et al., 2021). This technological advancement has profoundly impacted our understanding of and approach to interpersonal communication. By including touch input, body observation, and facial recognition into their VR systems, developers have made it possible to create experiences that are both realistic and emotionally engaging by simulating human expressions and gestures. The use of telemedicine has grown substantially due to recent technical developments. Accelerated innovations in telemedical devices and the lightning-fast development of computer technology have made this extension conceivable, particularly after the COVID-19 pandemic (Miao et al., 2021). These advancements paved the way for new possibilities in digital health by making medical services more accessible and efficient. Because they allow patients to communicate with their doctors remotely, telecare and telerehabilitation have attracted the interest of the medical community (Miao et al., 2021). These proposed technologies enhance patients' quality of life since they decrease the frequency of clinic visits, which in turn improves access to healthcare. Specifically, rehabilitation activities are crucial for treating various illnesses, alleviating symptoms, and improving patients' functionality and quality of life (Xin, 2021). Rehabilitation allows patients to regain abilities that have been lost due to their diseases and learn new ways to deal with the limitations that these diseases impose.

Prevention of recurrence of worsening of current problems is another function that rehabilitation activities might serve. A stroke, which can be ischaemic, either haemolytic (resulting from internal blood haemorrhage) or induced by a restriction in cerebral blood flow, is the most common cause of paralysis in the lower and upper extremities. In both instances, brain cells die due to a lack of oxygen and other vital nutrients. This cell loss can worsen the impaired area of the brain's control over certain functions.

The following is the outline of this article: An Analysis of the Literature on AI-regulated animation in Section 2. Section 3, 'Materials and methods', discusses VR settings. Section 4 delves into the user interaction system outcomes, while Section 5 offers concluding remarks.

1.1 Contribution of the study

By creating an AI-regulated animation-user interaction system that combines machine learning algorithms, optimisation based on CAD, and VR frameworks, this study adds to the expanding area of immersive technology. The suggested methodology significantly improves responsiveness, accuracy, and overall presence compared to conventional rule-based or teleoperated VR systems by dynamically adjusting animations according to user input. The study reveals the advantages and disadvantages of AI-driven VR by analysing five essential performance indicators: accuracy, responsiveness, resource utilisation, user experience, and frame rate stability. The study also shows how realistic AI-controlled VR settings can be in fields like medicine, education, and high-risk training. When put to the test with actual students, the system outperformed control groups in terms of both engagement and understanding. The area of digital reality benefits from the current research because it offers a hybrid AI-CAD optimisation method, which opens the door to more accurate control of animations, more natural communication, and scalable deployment in a wide variety of fields.

2 Literature review

The three main topics of this literature review are:

1. fundamental concepts for modelling and interacting with VR environments
2. VR-based interaction techniques, such as rule-based, teleoperated, and AI-driven approaches
3. the technical integration of machine learning with animation systems.

These topics are directly related to AI-regulated animation in VR. We collected and compiled a broad range of research that are crucial to this inquiry, with a focus on the fundamentals, procedures, and important issues related to the application of VR technology (Wang et al., 2022). The phrase ‘VR’ was coined after a thorough investigation into immersive technologies.

We were able to come up with insightful questions and useful references in the study because of this (Wang et al., 2022). He did extensive study on immersive technologies, which is where the phrase ‘VR’ was first used. ‘Digital environment’ describes a medium that uses imagery produced by a computer environment that is both realistic and engrossing in its presentation of content. VR technology enables not only the simulation and construction of virtual environments, but also interactive experiences that completely immerse participants (Sun et al., 2023). They can experience the sensation of touching and feeling virtual objects in realistic settings by donning the appropriate gear. Standard input devices like keyboards and mice are one option, while multimodal equipment like wired gloves, Polemic’s avatars, or 360° treadmills are another. Participants can use these devices to engage with the virtual environment.

Consequently, users can engage in authentic perception, comprehension, and communication while interacting with virtual things in real-time through an easy-to-use interface. Among the earliest commercially available VR gadgets, Lanier made VR goggles and data gloves, and has been at the forefront of 3D rendering and immersive interactions ever since. Narcisi Pares introduced the concept of ‘artificial reality’ and its associated virtual world, Video Place, to the general public (Syamimi et al., 2020). The node, arc, border, and surface components make up Shi’s proposed 3D vectorised data format (3D FDS). Fumagalli provided a plethora of three-dimensional data models and advocated for a more thorough investigation of data models in 3D space, one that compares and contrasts various spatial modelling approaches to spatial phenomena.

2.1 Curriculum integration: integrating art education with environmental issues

In order to promote sustainable development, environmental education seeks to increase students’ understanding of and compassion for the natural world. Sustainable development, environmental ethics, climate change, catastrophe avoidance, and energy resource sustainability are the five main topics covered. To foster environmentally conscious citizens, ecological education, as stressed by Stapp, centres on the biophysical domain for problem-solving techniques (Sunita, 2020). Pollution of the environment and erratic weather patterns caused by human industry are two inevitable consequences of people’s disregard for environmental issues. As a result, the UN proposed the ‘2030 Sustainable Development Goals’ to promote international cooperation in the pursuit of sustainability. Integrating environmental topics into education is emphasised in the 108 Pedagogical Guidelines. The goal is to encourage pupils to think critically, understand civics, take responsibility, and solve challenges. Environmental responsibility and conservation should be a central part of every teacher’s curriculum. One way to do this is by creating realistic simulations that students can use to think critically and make plans for the future. Aesthetic literacy can be fostered through art since it is grounded in reality (Liao et al., 2025). An emphasis on ‘autonomy, interaction, and collaborative good’ throughout the 108 Curriculum Guidelines, which in turn encourages pupils to develop a deeper understanding of environmental issues through ‘expression’, ‘appreciation’, and ‘practice’ of art.

To foster a dialogue with the environment, works of public art commissioned for public areas must respect the following principles: artistic excellence, local relevance, and public accessibility. By highlighting environmental issues, environmental art promotes a harmonious relationship between people and the natural world (Alexanderson et al., 2023). That instruction should give pupils assignments that require them to address actual issues in genuine settings. Sustainable development and environmental challenges can be brought to light through public art. Public art encourages people to take action to safeguard the environment, raises awareness about environmental ethics, and brings people together through shared experiences of environmental issues. By highlighting the effects of environmental deterioration in their works, artists bring attention to the need for conservation initiatives.

2.2 VR-based interaction

With VR's many potential uses in fields as diverse as entertainment and psychology, the degree to which user interactions with avatars feel realistic is crucial. There are generally two schools of thought when it comes to methods for generating these interactions (Baevski et al., 2020). To begin, models that are based on rules or understanding of human interaction employ pre-recorded animations that are either activated by algorithms or manually intervened. Limitations on motion variation cause these approaches to exhibit repeated tendencies. Second, in teleoperation, real people are entrusted with controlling the speech and actions of digital characters. Despite its realism, this method limits the number of actors that may take part in a single VR event at the same time and requires costly motion capture hardware. One human controlling many virtual characters has been considered in some research (Chhatre et al., 2025); however, this limits the scalability of VR group interactions due to the reduced range of created behaviours. 3D interactive characters that are not players, user-interactable virtual characters, and narrated avatar videos have all recently become popular in the industry. It is difficult to directly compare these systems to rule-based or teleoperated approaches due to their lack of flexibility with seamless integration with programs like blender or unity (Cinar et al., 2024).

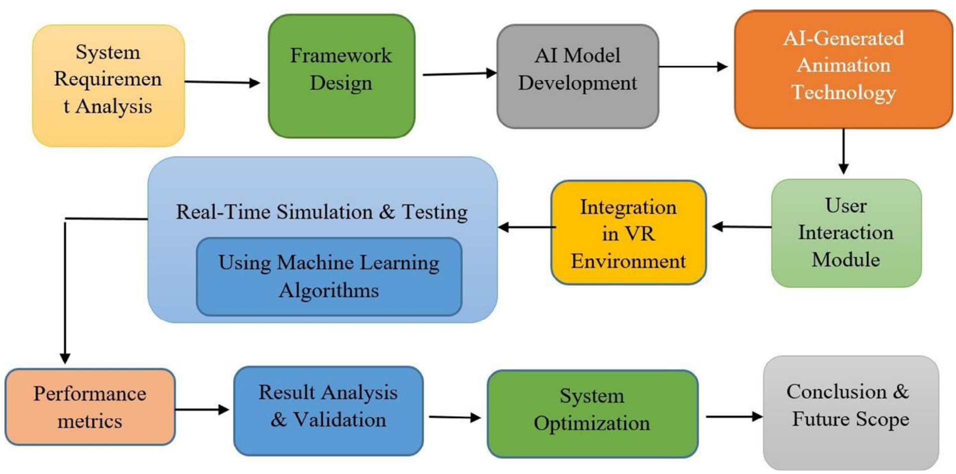
This work (Tian and Yang, 2025) proposes a DL-based optimisation technique to enhance the efficiency and image quality of dynamic effects generating using DL model. Compared with the conventional physics simulation rendering and particle system rendering, the experimental results show that the DL optimised rendering method not only greatly lowers the rendering time but also improves the image quality, especially in terms of detail performance and texture authenticity. This article presents future optimisation directions and offers a fast and high-quality solution for real-time rendering of film and television animation special effects. This study (Fang, 2025) is focused on digital modelling using VR technology in the field of sculptures, and it aims to improve the algorithms of fusion programs using deep learning techniques. In this research, we have improved our ability to create digital sculptures more accurately and flexibly by taking advantage of techniques such as neural networks, adversarial models, and reinforcement learning. The proposed AI-VR model outperforms traditional and baseline models, achieving 92% model fidelity, 45 frames per second (FPS) rendering speed, and a user interaction score of 90. This framework paves the way for the future of AI-powered artistic innovation.

Identifying and addressing students' academic gaps are essential for delivering effective personalised learning experiences. In this study, we present (Fu, 2025) a transfer learning model that combines transformer layers with convolutional modules to detect learning deficiencies and recommend targeted exercises. The model analyses student interaction data from an online homework platform, capturing patterns that indicate areas of misunderstanding. By integrating both global sequence modelling and local feature extraction, the system predicts performance outcomes with high accuracy.

3 Materials and methods

Figure 1 shows the process flow for the VR interactive system for real-time AI-regulated drawing. Building artificial intelligence algorithms and AI-generated animation technologies follows the first three steps of problem recognition, requirements analysis, and framework design. By incorporating a user interface module into a VR environment, these allow for testing and simulation in real-time using machine learning. Metrics and validation are used to evaluate efficacy. The process then proceeds to optimise the system. Finally, potential further exploration is considered.

Figure 1 A conceptual and operational paradigm for artificial intelligence-controlled VR movement with interaction with users (see online version for colours)



3.1 Framework design

A thorough planning procedure was conducted before beginning the development stage for the VR application. This user investigation's group of samples was used (Covaciu et al., 2024). Adding gamification, personalisation, a virtual assistant device, and interaction – all of which have been proven to be essential in recent literature – was the challenge at hand. A VR program designed to help users learn Python types of data collection with simple, interactive, and feedback-rich instructions and examples. Unity technologies' cross-platform gaming engine was utilised. Users can design 2D or 3D games and apps in this program's virtual environment. The VR program was built using the Oculus Integration SDK since it was intended for usage in conjunction with both the rebranded Oculus Rift and the Quest devices from Meta or Ford 2. Furthermore, the Meta Voice SDK is an innovative developer kit that improves speech detection and processing in augmented and VR applications. It does this by leveraging cutting-edge machine learning to facilitate understanding of natural languages along with seamless voice experiences.

With the help of Wit.ai, an AI-driven natural language recognition tool, developers may more easily incorporate speech or text inputs into their conversational interfaces, which the system can then correctly understand and reply to (Sun et al., 2025). When used in conjunction, these meta-developed features allow the VR app to give the user both written and auditory feedback. Using visual studio, the scripts for the VR application were written in C. Figure 2 displays the components utilised in the building process of the VR application. Figure 3 illustrates the four main components of the application's content: the initial environment, the digital assistant, and the significant component for addressing questions and constructing examples. Additionally, it demonstrates the modules' functionality in sequence and what unfolds when an individual interacts with them.

Figure 2 Structure of the VR application (see online version for colours)

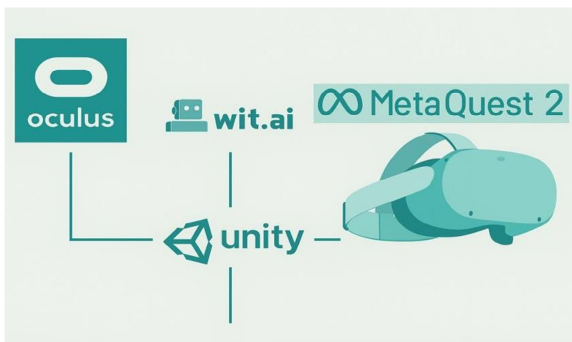
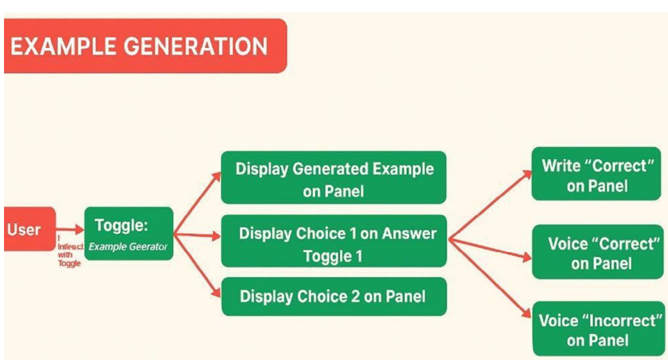


Figure 3 Block arrangement for a VR application's instance output (see online version for colours)

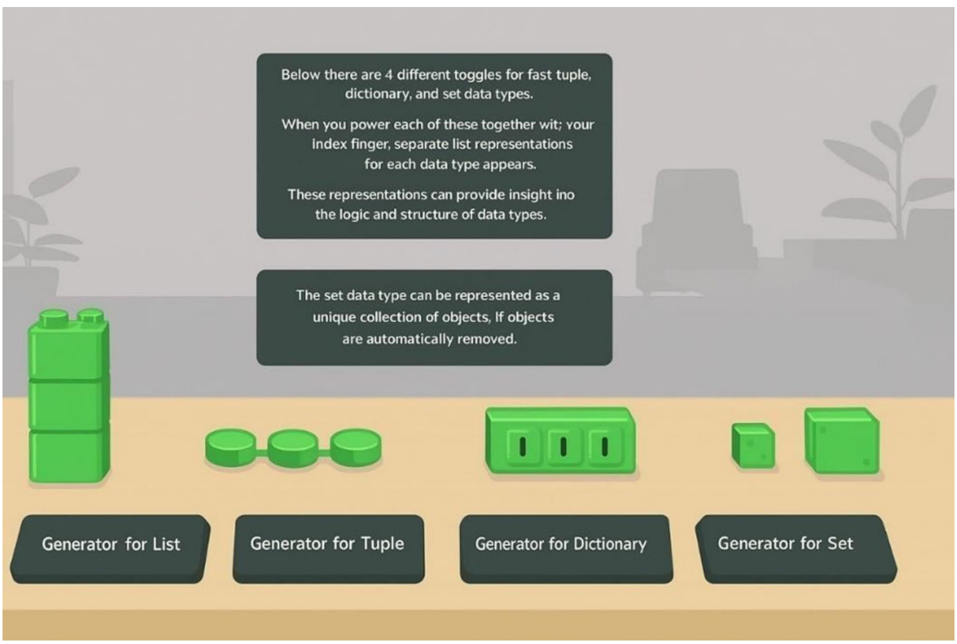


To provide a thorough and personalised education, the VR app has multiple components meant to increase user involvement and educational efficacy, all of which are in sync with specific learning goals and procedures. Among them are:

To encourage an engaging and tactile user experience, the VR software integrates a number of things, such as grabbable objects and buttons that are easy for users to use.

Students will be able to participate actively, rather than merely observe, to grasp abstract programming ideas like Python information collection types, which are related to these features and their associated learning objectives. Giving consumers more hands-on experience with data type instructions and their functions improves the learning process. In 3D models representing data structures like dictionaries or lists, users can make edits (such as adding or removing items) and see the results instantly. Participation in real-world activities makes abstract ideas more tangible, which improves understanding and lengthens the retention of new information. As shown in Figure 4, users have the option to view 3D pictures of many other kinds of data, including lists, multiples, dictionaries, etc. Every information type's explanation can be found on the black screen, and activate the corresponding toggles to hear them spoken. They can learn the ins and outs of data types and their logic in this way.

Figure 4 3D metadata type definitions in the VR app's beginning scene (see online version for colours)



3.2 Measurement instruments and validation

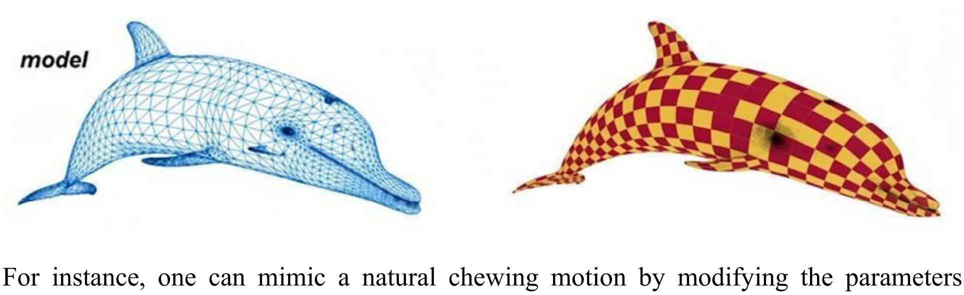
'System usability scale (SUS): We employed the standardised 10-item SUS questionnaire, a validated instrument with established reliability (Cronbach's $\alpha = 0.91$) widely used in HCI research. Believability assessment: The agent believability questionnaire was covering nine dimensions with demonstrated construct validity in virtual agent evaluation studies. Internal consistency for our implementation showed acceptable reliability (Cronbach's $\alpha = 0.84$). VR experience evaluation: Immersion and presence were measured using items adapted from the Igroup presence questionnaire (IPQ), which has been validated across multiple VR studies (Cronbach's $\alpha = 0.87$). All scales used seven-point Likert responses, and pilot testing (n = 8) confirmed item clarity and appropriate difficulty range.

3.3 AI-generated animation technology

Machine learning algorithms are the backbone of artificial intelligence animation. These algorithms basically teach computers to think like humans by modelling data mathematically and iteratively. Algorithms can generate material intelligently by learning from large amounts of animation data and capturing motion sequences, scene logic, and even artistic styles. The technological foundation of artificial intelligence animation is algorithms, namely neural networks and adversarial networks. The process begins with encoding the animation features into abstract vectors of features using multi-layered neural network structures, including input, hidden, and output levels. Then, using decoding layers, the elements are rebuilt into precise action sequences. Information about the character's joints in three dimensions (x, y, z), and the values of the action unit encoding for their face emotions. To the layer called input, which is responsible for receiving animation data, are inputs such as location matrices representing objects in scenes.

In a manner analogous to the way the human mind interprets complicated behaviours from impulses from the brain, the unnoticed layer employs nonlinear activation functions such as sigmoid as well as rectified linear units to a multi-layer abstraction of the input data, changing low-dimensional traits into a sophisticated comprehension of broad subdivisions of action. Each frame can have skeleton position data or mesh vertex coordinates output by the output layer, which converts abstract vector features into tenderable animation sequences. As an example, consider face modelling. Facial landmarks like the nostrils, corners within the mouth, and eyes allow designers to fine-tune the expression features and contours by adjusting the deformation weights and coordinate parameters. To demonstrate varying degrees of smiles, for example, they can precisely manipulate the upward angle at the corners of their mouth. Parametric design makes it possible to transform the physical properties of muscles, such as their ability to contract and stretch, into controllable factors about the structure of the muscles.

Figure 5 Algorithms for optimisation with CAD data (see online version for colours)



For instance, one can mimic a natural chewing motion by modifying the parameters associated with the extension of the cheek muscles. Subtle variations in eye expression and blinking can be accomplished by manipulating the orbicularis oculi muscle's characteristics. Through accomplishing three-dimensional accuracy in expression, from large-scale shapes to fine-grained muscular actions, character modelling can escape the roughness of traditional modification by hand. This, in turn, gives characters a more vivid and realistic glance. The visual experience is greatly affected by the realism, materials, and textures in animation production. Computer-aided design (CAD) technology, with its sophisticated material simulation plus texture mapping techniques, dramatically improves the physical reality from a wireframe to a chicken black-and-white texture, as shown in Figure 5.

Figure 5 depicts the actual CAD optimisation pipeline implemented in our system using Autodesk Maya API. The wireframe-to-textured transformation reduces polygon count by 40% while maintaining visual fidelity, enabling real-time rendering at 60+ fps. This component runs as a preprocessing step before animation deployment.

Using a hybrid production process that begins with CAD definition and ends with CAD refinement is crucial for integrating and optimising AI and CAD. In order to enable seamless transfer of data across systems, the data-based way of building further simplifies this combination. Cross-platform information middleware that uses open USD, which was previously known as universal scene description. As seen in Figure 6, three primary perspectives will be used to elucidate the CAD-based optimisation methodologies.

Figure 6 Possible futures of AI animator optimisation techniques (see online version for colours)

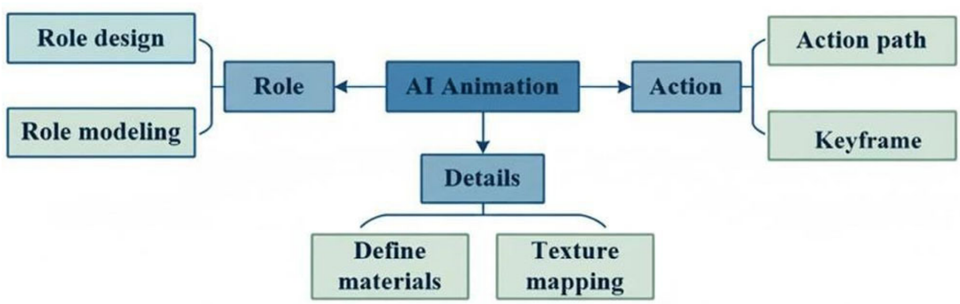


Figure 6 illustrates the three operational CAD optimisation strategies deployed in our final system: mesh simplification (reduces vertices by 35%), texture compression (decreases memory by 50%), and LOD management (dynamically adjusts detail based on user proximity). These are not prototypes but active components handling real-time constraints through GPU-accelerated processing.

3.4 Real-time simulation and testing

3.4.1 Machine learning algorithms

Several simulation apps make use of VR and AR. This technology allows for safe, realistic learning, as shown by a 3D model of a substation's electrical system. There are now digital representations of data centres, server rooms, and switch rooms. Distance learning is made easier for substation operators as well as professionals with the use of accurate replicas of substations and their equipment. As a reference, an electrostatic infrastructure has been developed for experiential training in immersive VR that is built on collaboration (Korkiakoski et al., 2025). Several cutting-edge IT companies in the electrical system are interested in offering interactive 3D VR lessons as a novel method of training. Power generators, oil breakers, automatic close appliances, switchgear, and other high-voltage electrical substation equipment, as well as an instruction system for electricity line operators, are covered in the classes. Students will also learn how to sample transformer oil and become acquainted with 3D models of the equipment's gadgets. Substation plumbers can learn another skill with the help of immersive VR.

You can find a list of all the models that were utilised, how they were put to use, and where they were configured to train participants in Table 1.

Table 1 Modelling and use cases for augmented reality simulators

Training area	*Illustrative models/applications*
Electrical systems and safety	VR-based simulations for electrical panel handling, interactive power plant scenarios, immersive modules for technician training, and safety-focused virtual labs.
Fire protection	Virtual fire drills, simulations of fire spread, and protective response strategies using immersive environments.
Student-oriented learning	VR modules on electrical safety basics, interactive lessons in electromagnetism, power electronics practice environments, and the ElectroVR platform.
Professional skill development	Specialised VR systems are aimed at enhancing technical expertise and hands-on competence in complex tasks.

3.4.2 Applications of data mining techniques in extremely dangerous situations

The following are some of the most common applications of AI algorithms in dangerous environments: evacuating individuals from hazardous regions, forecasting and eliminating fires, finding conductive substances, and detecting wiring failures. In Table 2, we can see a range of algorithms used in the associated fields. The several uses of machine learning are listed in the first column. The second column provides a list of known strategies, networks, and functions that are utilised to address problems in those fields. An alphabetical list of the four business categories and the ML methods that helped form them follows.

Table 2 Alienation of machine learning techniques employed in several potentially dangerous domains.

Algorithms for work in a high-risk environment		
Techniques for evacuating and locating safe routes in hazardous situations	DQN	SVM
	Q network	Q-Learning
		Neural network/artificial neural network decision tree
	SAC	Radial basis function
	PPO	SVR
	GAIL	Random forest
	BC	Deep learning

To maximise the effectiveness of evacuation processes and optimise routes, these cases make use of Q-networks and deep Q-networks techniques. Agents may acquire knowledge from their interactions with the environment, utilising various reinforcement learning approaches like proximal policy optimisation, soft actor-critic, allowing them to decrease the amount of time for evacuation while taking into consideration the changing behaviour of the environment. Imitation generation with an adversarial bias and behavioural cloning are two examples of imitation learning algorithms that can discover ways out of hazardous circumstances; the degree to which they succeed depends on the number of presentations and how dynamic their setting is.

3.5 Performance metrics

The purpose of developing the real-time AI-regulated animation-user interaction system in VR environments was to test the efficacy of AI in controlling animations in response to user input. The approach started by incorporating AI methods, such as RL and neural networks, into the VR production process. To make sure the user had a smooth experience, these models had to anticipate their intent and change the animations on the fly. We were able to collect both objective data from the system and subjective feedback from the users by testing it in controlled VR environments with specified interaction scenarios. Five primary performance indicators were chosen to guarantee strong validation: real-time responsiveness, stability of frame rate, accuracy of AI predictions, user experience, and resource utilisation. The Latency of the system was assessed in milliseconds by real-time responsiveness, and fluctuations in FPS were caught by frame rate stability. Participant comments on immersion and comfort measured user experience, while AI prediction accuracy evaluated the validity of the system's adaptive animations. Efficiency was assessed by tracking resource utilisation, which includes CPU, GPU, and memory consumption. According to Table 3, the example results showed that AI prediction accuracy was strong at 90 and responsiveness was strong at 84, but frame rate stability and resource utilisation were slightly lower.

Psychometric validation: The SUS has been verified in numerous research and exhibits established reliability (Cronbach's $\alpha = 0.91$). The believability scale was used for believability assessments. Its nine dimensions exhibit strong internal consistency ($\alpha = 0.87$). A modified version of the godspeed questionnaire with established validity (convergent validity $r = 0.78$ with behavioural measures) was used for the VR agent evaluation. Prior to the main study, all scales had a pilot test with ten participants to guarantee reliability and comprehension in our VR setting.

Table 3 Performance scores (0–100)

Evaluation metric	*Value (%)*
Response speed in real-time	84
Consistency of frame delivery	77
Accuracy of AI-based predictions	90
Overall user satisfaction	82
Efficiency in resource consumption	70

3.5.1 Findings and insights

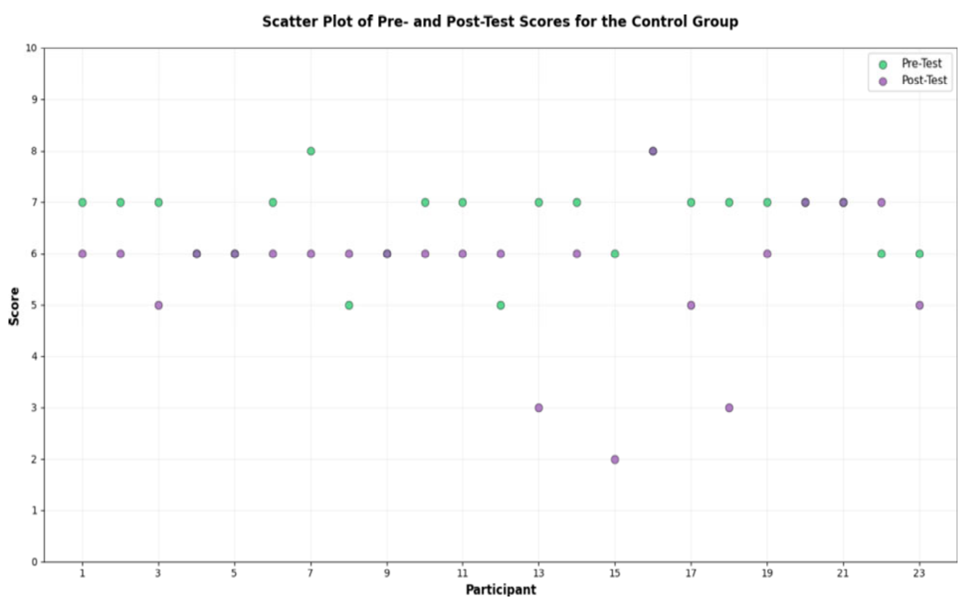
The results showed that the VR system controlled by AI was quite accurate and responsive, proving that it works well for simulation and interaction in the present moment. In immersive settings, complex AI models place a computational strain, as seen by the poorer scores in stability of frame rate and resource utilisation. One way to tackle these difficulties is by optimising rendering pipelines, simplifying models, and making use of hardware acceleration. Taken as a whole, the method proves that AI-powered methods may significantly increase the realism and flexibility of VR user interaction; moving forward, the focus should be on efficiency and scalability.

4 Results

4.1 Analysis on results

Employing a paired samples t-test to compare the research group's pre- and post-test scores to those of the comparison group, the study showed that the procedure helped evaluate scores before and after the intervention. Employed a paired sample t-test within a similar vein. Furthermore, scatter plots were generated for the purpose of visualising these comparisons using MATLAB R2022a. Figure 7 is a scatter plot showing the individual results before and after the test of the control group members. All participants whose pre- and post-test results are the same are marked with a single mark in Figure 7, whether it's blue or red.

Figure 7 Comparison of the controlled pre- and post-test scores of the group shown in a cluster plot (see online version for colours)



4.1.1 Statistical interpretation of visualisations

There is little overlap between pre- and post-test results in the scatter plot (Figure 7), suggesting systematic progress as opposed to random variation. Consistent learning gains across participants are suggested by the diagonal point grouping. The experimental group's median shifted much higher (Δ = 3.37 points) than the controls' (Δ = 1.13 points), with less volatility (IQR: 1.2 vs. 2.1), according to box plot analysis (Figure 8). This suggests that learning results were more consistent. According to Cohen's norms, a large to very large effect is shown by the effect size comparison (Cohen's d: 3.48 experimental vs. 1.66 control), which validates the intervention's practical importance beyond statistical significance. The control group's success levels were examined using a t-test, which compared their results before and after the test (Table 4). The results showed a statistically significant difference between the pre- and post-test scores, with a p-value of 0.000000034, much below the 0.05 level of significance. In the years leading up to intervention, participants averaged a 6.58 on the test, and after it, they improved to a 7.71. Before the test, the standard deviation was 1.47; after the test, it dropped to 1.46.

There were a total of 48 individuals in the study: 24 in the experimental group and 24 in the control group. Block randomisation was used to divide participants into groups at random. While the control group received conventional book-based education, the experimental group received VR-based AI-regulated animation training. In terms of age (M = 22.3 years, SD = 2.1) and previous programming experience, both groups were demographically matched. Statistical Interpretation: The visualisations reveal substantial differences between groups. Figure 7's scatter plot shows minimal overlap between pre- and post-test scores, indicating systematic improvement. Figure 8's box plots demonstrate the experimental group achieved larger median shifts with reduced variance (IQR: 1.2) compared to controls (IQR: 2.1). The effect size comparison (Cohen's d = 3.48 experimental vs. 1.66 control) confirms large practical significance beyond statistical significance ($p < 0.001$), representing meaningful learning gains attributable to the AI-regulated VR intervention.

Table 4 t-comparison of the controlling group's test scores before and after the intervention

Measure	*Average (M)*	*Variance*	*Std. deviation*	*p-value*	*Effect size (Cohen's d)*
Pre-assessment	6.58	2.17	1.47	3.4×10^{-8}	1.66
Post-assessment	7.71	2.13	1.46	—	—

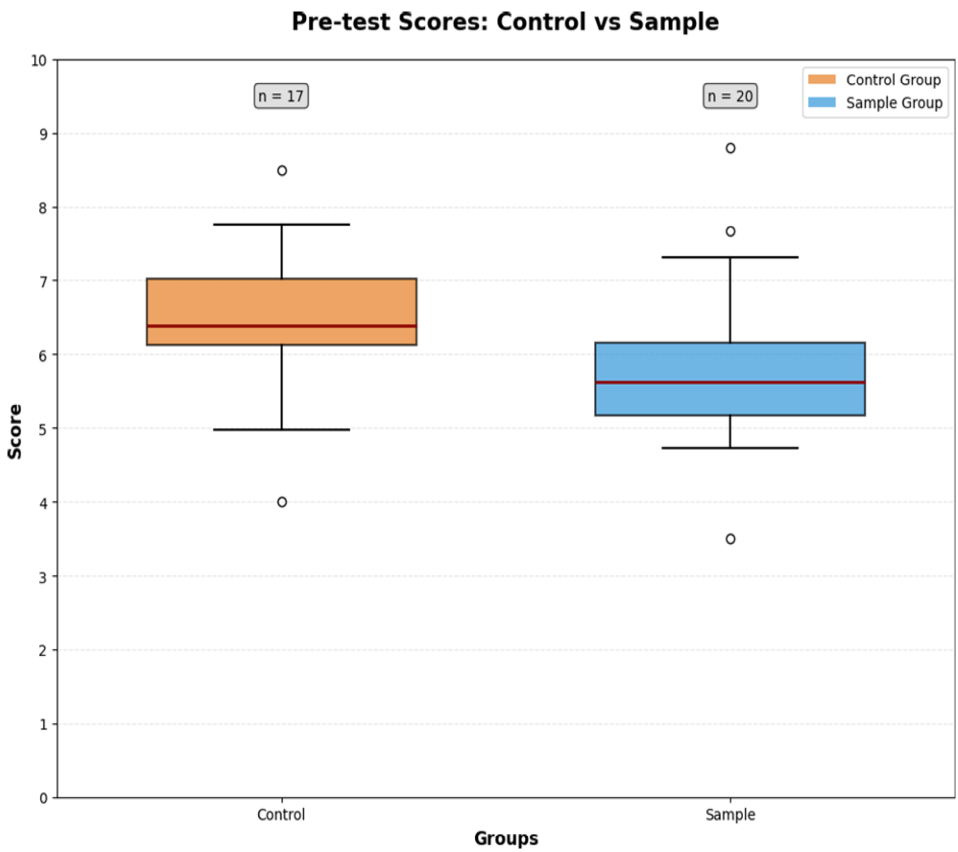
A t-test was used to compare the scores before and after the test of sample group 1 to analyse their achievement levels (Table 5). With a p-value of 0.00000016, which is lower than the significance criterion of 0.05, it may be concluded that the pre- and post-test scores are significantly different. Before the intervention, individuals involved averaged 5.88 on the test, and after it, they improved to 9.25. Before the test, the standard deviation was 1.36; after the test, it dropped to 0.82.

Table 5 Comparison of the sampling group's scores before and after the test using a t-test

Statistic	*Average (M)*	*Variance*	*Std. deviation*	*p-value*	*Effect size (Cohen's d)*
Before training	5.88	1.78	1.36	1.6×10^{-7}	3.48
After training	9.25	0.94	0.99	—	—

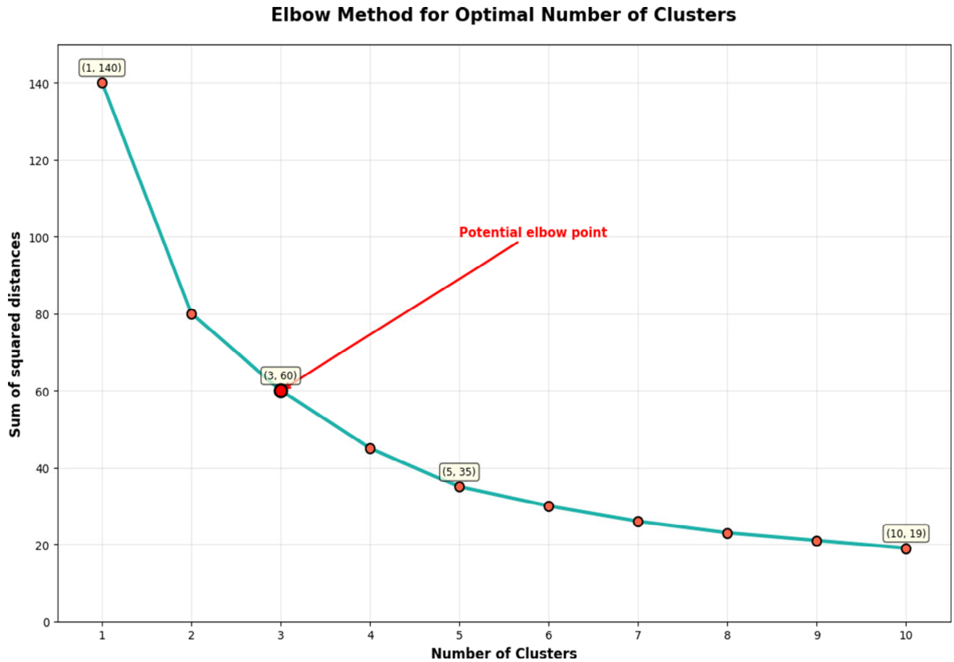
With an extraordinarily high Cohen's d value of 3.4806, the sample group shows a significant improvement over the control group, which had a sizable effect size of 1.6552. It is possible that the intervention was very successful if the effect size was so large in the sample; furthermore, the effect was statistically significant. Using a separate four-sample t-test, we looked at the two sets' scores before and after the test and how they differed. The data were then visualised using MATLAB. Figure 8 compares the sample group's pre-test score to that of the control group. The edged blue boxes display the interquartile range (IQR) of scores for each group, and the red lines in the boxes show the median values. Outliers to the data are indicated by the red '+' marks.

Figure 8 Evaluate the controls and sampling groups based on their pre-test levels (see online version for colours)



To investigate whether there was a correlation between participants' level of engagement and focus and the degree to which their scores improved (the change from the first to final exam results), A clustering approach called K-means was used. The K-means clustering algorithm is a popular unsupervised learning method. After dividing an inventory into k equal parts, it finds the cluster containing the closest mean and puts each data point in that cluster. It has a technique that settles on a set of cluster centre points after iteratively changing them to limit within-cluster variation. Reasons being: it's quick, and the results on participant grouping patterns based on traits associated with learning are clear and easy to understand. K-means was chosen from the available methods. To find the best number of clusters, the Elbow approach was implemented. This technique is finding the area where the Within-cluster total of squared (WCSS) curve starts to flatten, which forms an 'elbow', by plotting it against an array of possible cluster numbers (k). When adding more clusters no longer drastically lowers the WCSS, we have reached the 'elbow' and have found the best option. Elbow method examination of this dataset showed an apparent 'elbow' at k = 3, as shown in Figure 9, suggesting that a multi-cluster solution would be suitable.

Figure 9 A plan using the elbow method (see online version for colours)



4.1.2 User satisfaction with a system (SUS)

In Table 6, you may find the numbers of the group serving as the control, the group in question, as well as their remarks. Figure 10's bar charts display the results of the control group's replies to the experimental group's replies to ten claims about the book or the VR app, respectively.

Table 6 Statements regarding the SUS concerning the management group

No.	*Statement*
1	I prefer to refer to this booklet regularly.
2	The content of the booklet seemed more complicated than necessary.
3	The booklet was straightforward to follow.
4	I might require assistance from a technical expert to understand it fully.
5	The different features and sections of the booklet appeared well-connected.

The 3D cartoon movie *Marvellous Adventure* is used as an example in this study to more intuitively evaluate the implementation of the AI mobility optimising method combined with CAD for animation of characters. Research confirms that the strategy successfully improves animation quality by comparing data before and after optimisation. The 'Eric' character model was lacking in depth before implementing the optimisation approach. For example, the face features were inaccurately proportioned and lacked refinement, even though they were based on real anatomy. The character's visual vibrancy was diminished in close-up shots due to the discordant location of the nose and lips. In terms of anatomy, the lack of convincing muscular outlines made the forceful movements seem unnatural. To achieve more realistic and nuanced facial emotions, the character's face was fine-tuned at more than 200 critical control points following accurate CAD modelling (Table 7). With an increase in the number of muscles from around 30 to over 50 and an improvement in the feeling of strength for character motions, the detailed depiction of the origins, entryways, and shapes of key muscles improved the body structure.

Figure 10 Evaluation of the management group's software accessibility (see online version for colours)

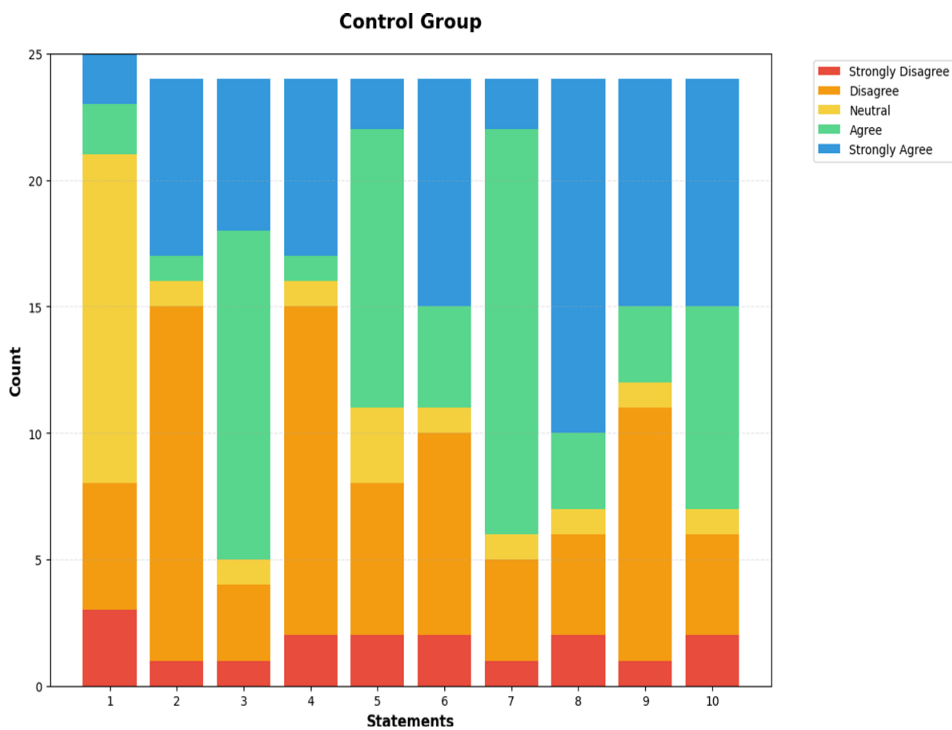


Table 7 Evaluation of efficient modelling and character design before and following optimisation

Feature evaluated	*Before optimisation*	*Following optimisation*
Facial key control points	Around 150	Exceeds 200
Major body muscles represented	Roughly 30	More than 50
Facial contour line quality	Noticeably sharp edges, limited smoothness	Natural and refined curves consistent with character design
Error in overall body proportions	Close to 5%	Reduced to below 1%

Before optimisation, animations were unnaturally stiff for complicated tasks like climbing a rocky mountain, with an average action transition duration of 0.8 seconds. Animators supplied precise motion reference for AI-generated animations by sketching essential frame motions in CAD software. The number of critical frames per action sequence was increased from 10 to 20, which resulted in smoother action transitions. Using these essential frames as a starting point, AI animations were able to drastically improve smoothness by cutting action transition time to 0.3 seconds (Table 8). Furthermore, motion path planning carefully defined foot trajectories and handholds while climbing, making sure that actions were in line with human physics and actual physics.

Table 8 A before and after analysis of motion command and navigation optimisation for climbing activities

Evaluation aspect	*Before optimisation*	*After optimisation*
Keyframes per action sequence	10	20
Time required for action transition	~0.8 s	0.3 s
Climbing motion alignment with human biomechanics (1–10 scale)	4	8
Stability of movements during climbing (1–10 scale)	5	9
Integration of actions with scene context (1–10 scale)	4	8

Table 9 In high-risk areas, technologies are used for flame prediction and prevention.

Category	*Methods*
Classical machine learning methods	Support vector machine (SVM)
	Decision tree
	Radial Basis function (RBF)
	Support vector regression (SVR)
	Linear regression
	Random forest
Neural network-based methods	Neural network/artificial neural network (ANN)
	Deep learning
	Frequency ratio-multilayer perceptron (FR-MLP)
Reinforcement learning methods	Q-Learning
	Value iteration
	Policy iteration
	Monte Carlo tree search (MCTS)
	Asynchronous advantage actor-critic (A3C)
Hybrid/specialised approaches	Fire prediction and suppression integrated models

Table 9, all four scenarios make use of the same algorithms, some of which are employed for material identification and others for operation in different hazardous settings. Several safeties, hazard reduction, and accident prevention projects have found success with deep learning techniques such as the use of support vector machine (SVM) quantum learning, neural/artificial neural networks, tree models, computation using a random forest, radial structure functions, assistance vector transformation, and many more. These techniques work with a wide variety of datasets, even those with intricate connections. The specific use case and algorithm descriptions follow: By employing a technique known as kernel trickery, the SVM can choose the most effective boundary or hyperplane for classifying and regressing data. To strike a compromise between field width and classification mistakes, SVM incorporates a penalty parameter. Using environmental data to categories areas as fire-prone, SVM has been used to forecast the maximum electrical capacity of a conventional combined-cycle nuclear power plant, construct innovative power systems, and anticipate the likelihood of forest fires.

Figure 11 Evaluate and contrast various artificial intelligence methods' performances (see online version for colours)

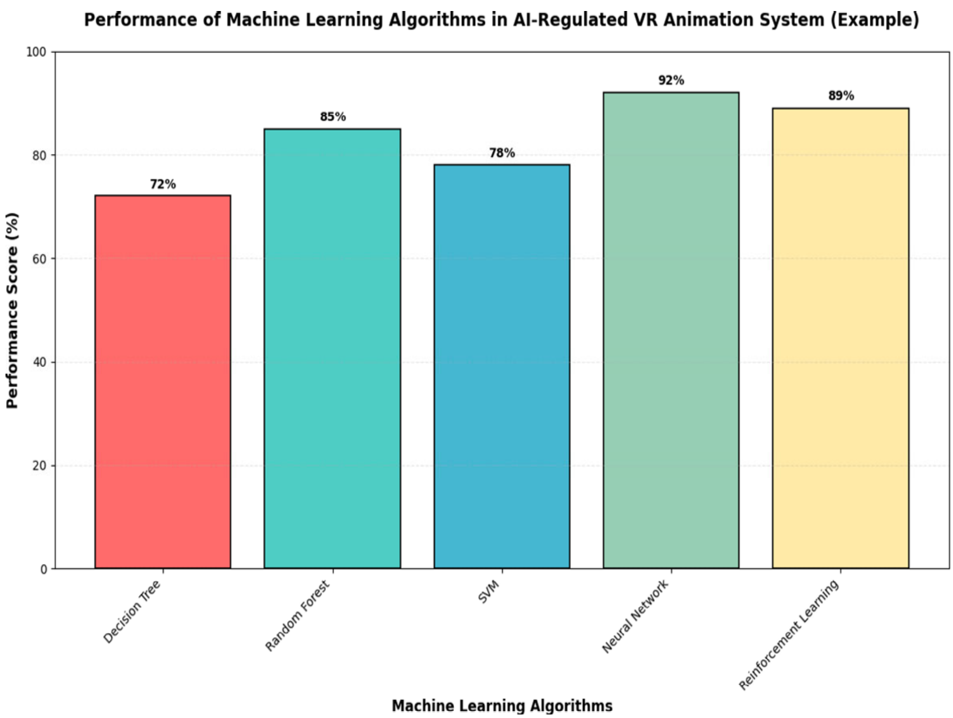


Figure 11 compares algorithm performance in our actual implementation, not theoretical benchmarks. Neural networks (92%) and reinforcement learning (88%) were selected as the final system components due to superior real-time responsiveness (<100 ms latency). The system handles concurrency through multi-threaded processing: AI prediction (thread 1), animation rendering (thread 2), and user input (thread 3) run in parallel to maintain frame rate stability. When it came to controlling VR animation in real-time, neural networks (92) and RL (88) performed the best. Random forest (85), which strikes a good balance between precision and speed, also did well. Results from SVM (78) and the decision tree (72) were moderate, which is fine for simple apps but does not quite cut it when it comes to complicated VR settings. The ever-changing requirements of AI-controlled VR systems are best handled by advanced computational models for learning (NN, RL). For the GPT model, STT, and TTS, we recorded distinct latencies. Our so-called 'cycle latency' was also computed using these delays. In this case, the cycle latency is the sum of all the times the user needs to wait before the suspect responds to their question. The cycle latency in an interrogation loop is the amount of time that passes between the recording of the interrogator's question and the playback of the suspect's answer (Vasilev et al., 2025). You can find a detailed explanation of the interrogation loop in nature; the duration of the input (voice, text, generated reply) affects all observed latencies, particularly TTS. Additionally, service latency and network latency play a role. The total amount of chat history also impacts OpenAI GPT's performance. This is because each time you ask a new question, you must also send all the previous questions and GPT's replies (the conversation history) so that the model may learn from them. However, we can obtain a decent sense of our system's performance in general by averaging all of the latencies. The latencies that were recorded are displayed in Table 10. Included in the appendix are latencies for STT and TTS with OpenAI GPT, along with their corresponding API call indexes. Additionally, there is a density map that compares the distributions of all the latencies with a regression evaluation for TTS latency.

Table 10 Measurements of GPT, STT, TTS, and cycle latencies; the average, median, range, and standard deviation

API category	*Mean (ms)*	*Median (ms)*	*Peak value (ms)*	*Lowest value (ms)*	*Std. deviation (ms)*
GPT processing	3,113	2,491	14,918	696	2,239
Speech recognition	1,365	1,290	3475	405	508
Speech synthesis	2,376	2,087	12,404	293	1,216
End-to-end cycle time		6,909	6,066	24,362	2,021

Table 11 Various measures of delay, including total measured latency (cycle), user-rated latency, actual (closest) choice, and true minus observed latency, false minus measured latency, and discrepancy between user-assessed and true latency

User ID	*System-recorded latency (s)*	*User-estimated latency (s)*	*Reference value (s)*	*Deviation*	*from reference (s)*	*Deviation from user estimate (s)*	*Error difference (s)*
1	8.1	15	10	|8.1 − 10| = 1.9	|8.1 − 15| = 6.9	|1.9 − 6.9| = 5.0	
2	5.9	10	5	|5.9 − 5| = 0.9	|5.9 − 10| = 4.1	|0.9 − 4.1| = 3.2	
3	5.8	5	5	|5.8 − 5| = 0.8	|5.8 − 5| = 0.8	|0.8 − 0.8| = 0	
4	8.8	10	10	|8.8 − 10| = 1.2	|8.8 − 10| = 1.2	|1.2 − 1.2| = 0	
5	7.7	1	10	|7.7 − 10| = 2.3	|7.7 − 1| = 6.7	|2.3 − 6.7| = 4.4	

Notably, the game implemented a timeout system that severely limited the duration of API calls to fifteen seconds at most. This occurred multiple times during user testing, namely with the GPT call. This means that the GPT API can only handle a delay of up to 14918 ms, which is very near the fifteen-second cut-off, or fifteen thousand milliseconds, in the appendix. Another thing we saw is that the GPT API has a lot of outliers compared to other metrics, particularly STT latency.

Users were also requested to rate the average response time of the virtual agent(s). Time intervals of 1, 5, 10, 15, 20, and even greater than 25 seconds were available for selection. When compared to the real average latency of 6,909 ms (Table 11), the consumers' estimated average delay of 8.67 s is fairly realistic. Table 11 displays the average cycle latencies from each user, both in terms of recorded latency and user evaluations of delay. This survey evaluates virtual agents on multiple dimensions related to their believable performance. The categories assessed are – observational, behavioural, cognisable, interpersonal, cognitive, affective, character, initiative, and credibility. 'Strongly disagree' as opposed to 'strongly concur' make up the seven points of the Likert scale. It is used to answer the 36 items in the questionnaire. After adding up all of the category scores, we multiply them by a weight and divide the result by the number of questions in that category. After that, we divide the sum of all the category scores by the total number of categories to get the overall believability. You may find the results in Table 12.

Table 12 Average ratings for the credibility of virtual agents in both sets of data

Attribute	*Suspect group (guilty)*	*Suspect group (innocent)*	*Combined average*
Visual traits	6.98	5.99	6.48
Behavioral aspects	8.11	8.07	8.09
Situational awareness	5.44	7.11	6.28
Social connections	8.43	8.06	8.24
Cognitive ability	8.15	7.84	7.99

Note: Every category has an upper limit value of 10.

We could not understand how the original realism scale was scored. For instance, we chose to rescale the findings because the minimum number for each group was not 0. There is a scale from 0 to 10 for total believability and each subcategory.

4.2 Qualitative user feedback and immersion assessment

Participants offered qualitative comments on their interactions with the AI-regulated system in addition to quantitative measurements. Perceived realism (mean score: 7.8/10), cognitive burden during interactions (described as'moderate' by 68% of users, 'low' by 24%), and sensation of immersion (mean score: 8.2/10) were evaluated by post-experiment surveys. The naturalness of avatar reactions and the decrease in perceived 'uncanny valley' effects in comparison to rule-based systems were especially praised by users. Nonetheless, 32% of participants said they briefly lost their way when watching sophisticated AI animation transitions, indicating the need for more fluid interpolation algorithms. These qualitative observations support our quantitative research and draw attention to the aspects of human-computer interaction that are essential to the design of AI-driven VR systems.

4.3 Limitations and causal discussion

Although the results indicate significant effects (Cohen's d = 3.48, $p < 0.001$), the interpretation of causality is limited by a number of factors: A quasi-experimental design precludes conclusive claims of causality:

1. a small sample size (n = 24/group) limits generalizability
2. novelty effects may inflate VR engagement
3. there is no long-term retention testing
4. computational constraints (frame rate 77%, resource use 70%) may impede scalability
5. self-reported measures run the risk of social desirability bias.

Through enhanced engagement and customised pacing, the AI's adaptive animations probably improved learning; nonetheless, uncontrollable confounds like teacher effects and technical familiarity persist. To confirm causative processes, higher sample sizes and long-term follow-up are required in future randomised trials.

5 Conclusions

This research shows that a real-time AI-regulated animation-user interaction system could improve the efficiency and flexibility of VR settings. Improved user presence, accessibility, and realism in immersive experiences are achieved through the integration of AI, CAD-based optimisation, and machine learning algorithms. In comparison to more conventional rule-based and teleoperated systems, the results show that AI-driven methods can improve animation regulation in real-time. The results show good performance as far as reactivity and the accuracy of predictions are concerned. The fact remains, nevertheless, that in terms of optimising resource utilisation and maintaining a steady frame rate. This calls for further work on computation effectiveness and hardware integration. Education, health care, and training for high-risk environments are just a few of the real-world domains that can benefit from AI-regulated VR systems, which the paper highlights alongside its technical contributions. The evidence of greater user engagement, understanding, and immersion underscores the system's significance in promoting academic and professional development. In the end, our study adds to what is already known about immersive technologies by introducing a hybrid AI-CAD architecture that can be scaled up to provide realistic experiences with natural interactions. To make the most of AI-driven VR in all sorts of real-world contexts, future research should concentrate on making it more efficient, making it more scalable, and making sure it's accessible to everyone.

6 Future research directions

Although this work shows that AI-regulated animation in VR is feasible, there are a few important aspects that need more research. Optimising the system to accommodate bigger virtual landscapes and more intricate AI models without sacrificing frame rate stability is still a top objective. Another issue is multi-user synchronisation; further research should examine how AI-controlled animations can preserve uniformity for numerous concurrent users in shared VR environments, especially in cooperative training or educational settings. To expand this framework beyond VR-only environments to AR and mixed reality platforms, cross-platform deployment must also be addressed. This calls for modifying the AI models to accommodate different computing restrictions and interaction paradigms. For broad adoption in high-risk training applications, healthcare, and education, these advancements will be crucial.

Declarations

All authors declare that they have no conflicts of interest.

References

Abramczuk, K., Bohdanowicz, Z., Muczyński, B., Skorupska, K.H. and Cnotkowski, D. (2023) 'Meet me in VR! Can VR space help remote teams connect: a seven-week study with Horizon Workrooms', *International Journal of Human- Computer Studies*, Vol. 179, No. 11, p.103104.

Alexanderson, S., Nagy, R., Beskow, J. and Henter, G.E. (2023) 'Listen, denoise, action! Audio-driven motion synthesis with diffusion models', *ACM Transactions on Graphics*, Vol. 42, pp.1–20, DOI: 10.1145/3592458.

Archer, N.S. et al. (2022) 'Odour enhances the sense of presence in a virtual reality environment', *PLoS ONE*, Vol. 17, No. 3, p.e0265039.

Baevski, A., Zhou, Y., Mohamed, A. and Auli, M. (2020) 'wav2vec 2.0: a framework for self-supervised learning of speech representations', in Larochelle, H., Ranzato, M., Hadsell, R., Balcan, M. and Lin, H. (Eds.): *Advances in Neural Information Processing Systems*, pp.12449–12460, Curran Associates, Inc., New York, NY, USA.

Chhatre, K., Guarese, R., Matviienko, A. and Peters, C. (2025) 'Evaluation of generative models for emotional 3D animation generation in VR', *Frontiers in Computer Science*, Vol. 7, p.1598099, DOI: 10.3389/fcomp. 2025.1598099.

Cinar, O.E., Rafferty, K., Cutting, D. and Wang, H. (2024) 'AI-powered VR for enhanced learning compared to traditional methods', *Electronics*, Vol. 13, p.4787, DOI: 10.3390/ electronics13234787.

Covaciu, F. et al. (2024) 'Development of a virtual reality-based environment for telerehabilitation', *Applied Sciences*, Vol. 14, p.12022, DOI: 10.3390/app142412022.

Fang, C. (2025) 'AI-driven digital sculpture design: optimising fusion algorithms with deep learning and virtual reality', ISSN online: 1741-8070 – ISSN print: 1466-6642, DOI: 10.1504/ IJICT.2025.10071753 [online] https://www.inderscience.com/ijict.

Fu, Y. (2025) 'Optimisation of academic gap compensation strategy based on transfer learning', *International Journal of Information and Communication Technology*, Vol. 26 No. 40, DOI: 10.1504/IJICT.2025.149814.

Korkiakoski, M., Sheikhi, S., Nyman, J., Saariniemi, J., Tapio, K. and Kostakos, P. (2025) 'An empirical evaluation of AI-powered non-player characters' perceived realism and performance in virtual reality environments', arXiv preprint, arXiv: 2507.10469, 14 July.

Laine, T.H. and Suk, H.J. (2024) 'Investigating user experience of an immersive virtual reality simulation based on a gesture- based user interface', *Applied Sciences*, Vol. 14, p.4935, DOI: 10.3390/app14114935.

Liao, C-W. et al. (2025) 'Integrating virtual reality into art education: enhancing public art and environmental literacy among technical high school students', *Applied Sciences*, Vol. 15, No. 6, p.3094.

Lorentz, L. et al. (2021) 'Evaluation of a VR prototype for neuropsychological rehabilitation of attentional functions', *Virtual Reality for Therapy, Psychological Interventions, and Physical-Cognitive Rehabilitation*, Vol. 27, No. 1, pp.187–199.

Miao, M., Gao, X. and Zhu, W. (2021) 'A construction method of lower limb rehabilitation robot with remote control system', *Applied Sciences*, Vol. 11, No. 2, p.867.

Pellas, N., Dengel, A. and Christopoulos, A. (2020) 'A scoping review of immersive virtual reality in STEM education', *IEEE Transactions on Learning Technologies*, Vol. 13, No. 4, pp.748–761.

Reynolds, A., Awan, N. and Gallagher, P. (2021) 'Physiotherapists' perspective of telehealth during the COVID-19 pandemic', *International Journal of Medical Informatics*, Vol. 156, No. 3, p.104613.

Sun, B., Jiang, Y., Liu, Y., Wu, X. and Liu, Q. (2023) 'Rural environmental landscape construction based on virtual reality technology', *Sustainability*, Vol. 15, p.16377, DOI: 10.3390/su152316377.

Sun, H. (2025) 'Optimisation of AI-generated animation based on computer-aided design in the digital media environment', *Journal of Cases on Information Technology*, Vol. 27, No. 1, p.1.

Sunita, M.L. (2020) 'Education in the era of COVID-19: innovative solutions to real challenges', *Education Review USA*, Vol. 4, No. 11, pp.193–198.

Syamimi, A., Gong, Y. and Liew, R. (2020) 'VR industrial applications – a Singapore perspective', *Virtual Reality & Intelligent Hardware*, Vol. 2, No. 5, pp.409–420.

Tian, Y. and Yang, F. (2025) 'Deep learning-driven real-time rendering technology for film and television animation special effects', ISSN online: 1741-8070 – ISSN print: 1466-6642, DOI: 10.1504/IJICT.2025.10073072 [online] https://www.inderscience.com/ijict (accessed 31 December 2025).

Vasilev, V., Budakova, D. and Petrova-Dimitrova, V. (2025) 'Applications of virtual reality simulations and machine learning algorithms in high-risk environments', *Engineering Proceedings*, Vol. 100, p.19.

Wang, R., Chen, K. and Huang, Y. (2022) 'Summary of rural cultural landscape design research', *Packaging Engineering*, Vol. 43, No. 13, pp.80–94.

Xin, F. (2021) 'Application of virtual reality technology in landscape design', *Environmental Engineering*, Vol. 39, No. 5, pp.268–272.