Journal of Contemporary Educational Research, 2025, Volume 9, Issue 4
http://ojs.bbwpublisher.com/index.php/JCER
ISSN Online: 2208-8474
ISSN Print: 2208-8466

Research on the Production of Micro-Courses for Educational Digital Humans based on AIGC Technology

Yongqing Lin[1], Chunhong Luo[1], Ling Qin[2]*

[1]Beijing Jiaotong Vocational Technical College, Beijing 102200, China
[2]Beijing University of Civil Engineering and Architecture, Beijing 102616, China

**Author to whom correspondence should be addressed.*

Copyright: © 2025 Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0), permitting distribution and reproduction in any medium, provided the original work is cited.

Abstract: As one of the latest research hotspots in the field of artificial intelligence, AI-generated content (AIGC) can create virtual teaching scenarios, such as those involving educational digital humans, through content creation. This integration of AI and education is a product that will profoundly change the form of online education and the learning experience. Currently, most courses for educational digital humans still rely on traditional educational models and face many problems in practice. Based on an overview of AIGC technology and its application in micro-course production, this paper analyzes the current status of micro-courses for educational digital humans. It proposes the application of AIGC technology and educational digital humans in micro-course production, namely, using AIGC technology to drive the transformation of micro-courses for educational digital humans towards intelligence, thereby improving education quality and efficiency.

Keywords: AIGC; Education; Digital humans; Micro-course production

Online publication: April 28, 2025

1. Introduction

With the rapid development of digital education, micro-courses, as a new type of teaching resource, have been widely used in the field of education due to their short and precise nature and clear themes. The continuous advancement of artificial intelligence technology has gradually penetrated the field of education, bringing new opportunities and changes to micro-course production. As an innovative application of AI-generated content (AIGC) technology in educational scenarios, educational digital humans are gradually changing the production mode and teaching experience of micro-courses. This study aims to deeply explore the application of educational digital humans based on AIGC technology in micro-course production, analyze their advantages and existing problems, and provide references for the construction of micro-course resources in higher vocational education.

2. Overview of AIGC technology and its application in micro-course production

2.1. Overview of AIGC technology

AIGC refers to the production method of generating content using artificial intelligence technology. It is based on various technologies such as deep learning, natural language processing, and computer vision, enabling machines to independently create text, images, audio, video, and other content. Its core principle lies in learning and training from large amounts of data to enable the model to grasp data features and patterns, thereby generating content with a certain degree of logic and creativity. The development of AIGC technology has gone through multiple stages. Early on, rule-driven artificial intelligence systems could only generate simple and patterned content. With the rise of deep learning algorithms, such as neural networks and transformer architectures, AIGC technology has made significant breakthroughs. Presently, advanced AIGC models such as GPT series and Midjourney continue to emerge, capable of generating high-quality and diversified content, which has been widely used in multiple fields.

2.2. Application of AIGC technology in micro-course production

2.2.1. Text generation

In the process of micro-course production, AIGC technology can be used to generate teaching scripts, knowledge point explanation copywriting, and more. Teachers only need to input information such as course themes, teaching objectives, and key knowledge points, and AIGC tools can quickly generate structurally complete and logically clear text content. For example, when producing a micro-course on historical events, after the teacher enters the theme "Battle of Red Cliffs," the AIGC tool can generate detailed explanation copywriting, including the background, process, and impact of the war. It can also generate text in different styles according to teaching needs, such as storytelling and academic analysis, saving teachers a lot of preparation time.

2.2.2. Voice generation

The voice generation function of AIGC can convert text into natural and smooth speech. By training on a large amount of voice data, the model can simulate different voice tones and intonations to generate realistic voice explanations. Teachers can choose voice types that suit the style of their courses, such as a gentle female voice for language courses or a steady male voice for history and philosophy courses. Additionally, the speech rate and intonation of the voice can be adjusted to enhance the infectiousness and attractiveness of the explanations. For example, in English listening micro-courses, the AIGC voice generation function can be used to generate standard English pronunciations, assisting students in their listening training.

2.2.3. Image generation

With the help of AIGC technology, corresponding images can be generated based on text descriptions, enriching the visual effects of micro-courses. Whether it is abstract concepts, complex experimental devices, historical scenes, or scenes from literary works, intuitive and vivid images can be created by inputting keywords and descriptive sentences into AIGC tools. For instance, when explaining the concept of magnetic fields in physics, AIGC can generate images showing the distribution of magnetic field lines, helping students better understand the abstract concept of magnetic fields.

2.2.4. Video generation

Some AIGC platforms have video generation capabilities, allowing them to integrate elements such as text,

images, and voice to automatically generate complete video content [1]. Through preset templates and animation effects, professional-level micro-course videos can be quickly produced. Teachers only need to provide key content to easily obtain micro-course videos with dynamic images and narrated explanations, greatly improving the efficiency of video production.

2.2.5. PPT generation

AIGC technology can generate PPT outlines and page layouts based on course content. After inputting the course theme and main points, the tool automatically matches appropriate templates to generate a PPT draft including a title page, content pages, summary pages, etc. The generated PPTs exhibit high aesthetic appeal and professionalism in terms of page layout and chart design [2]. Teachers can then make personalized modifications to perfect the PPT content.

2.2.6. Educational digital human generation

AIGC technology has the capability to create highly realistic educational digital humans, which can simulate teachers in delivering course explanations. Through deep learning, these digital humans can not only accurately convey knowledge but also match their expressions and movements with the content of their explanations, enhancing the vividness and interactivity of teaching [3]. Digital humans can be customized with different images and styles according to the needs of the course, such as a young and lively image suitable for primary school courses or a mature and steady image suitable for higher education courses.

3. Overview and production process of AI-generated digital human technology

3.1. Definition of educational digital human

An educational digital human is a virtual character created based on artificial intelligence technology, specifically applied in the field of education. It possesses multiple functions, including the ability to simulate human teachers in conducting teaching activities such as knowledge explanation and interactive question answering. By integrating technologies like speech recognition, natural language processing, and computer vision, the educational digital human achieves natural interaction with learners, providing them with personalized learning experiences [4]. It is also known as an “artificial human” or “virtual human,” digital virtual human, or educational virtual digital human.

3.2. Driving methods of educational digital humans

AI-generated digital humans are virtual characters created through artificial intelligence technology, with the core being the integrated application of technologies such as speech recognition, image processing, and natural language understanding. Currently, there are two main methods available in the market: one is based on real human motion capture technology, and the other is based on virtual human motion capture technology. The latter requires the use of motion capture equipment for real-time tracking, transmitting the actor’s movement data to a computer, which then simulates the corresponding actions.

3.2.1. Real human-driven

A real human-driven educational digital human involves collecting data such as movements, expressions, and speech from actual teachers. Through digital twin technology, a digital human corresponding to the real-world person is created. The teacher conducts the lesson in a specific motion capture equipment and recording

environment, and the digital human presents behavior consistent with the teacher in real-time or post-production [5]. The advantage of this driving method is that it can maximize the preservation of the teaching style and emotional expression of the real teacher, making the teaching process more authentic and natural. However, the disadvantage is that it requires higher equipment and environmental standards, resulting in relatively higher production costs.

3.2.2. Technology-driven

A technology-driven educational digital human is based on artificial intelligence algorithms. It autonomously generates movements, expressions, and speech through learning from a large amount of teaching data and language models. This type of digital human can automatically adjust its performance style according to different teaching content and scenarios. Its advantage lies in lower production costs and the ability to quickly generate digital humans with different styles. However, it may not be as effective as real human-driven digital humans in terms of emotional delivery and personalized teaching.

3.3. Construction of evaluation index system for educational digital humans

AI + education has been widely used in classroom teaching, extracurricular tutoring, online education, and other aspects. From early speech recognition systems to current face recognition systems, and now to the popular DeepSeek this year, AI + education is gradually changing learning methods and even the entire education industry. The evaluation index system for educational digital humans is shown in **Table 1**.

Table 1. Evaluation index system for educational digital humans

Evaluation metrics	Evaluation dimensions	Evaluation results
PPT functionality	Template richness	Rich, covering templates for various subjects and teaching scenarios
	Editing convenience	Simple operation, supporting fast online editing and content replacement
	Interactivity with digital humans	Good interactivity, digital humans can naturally integrate into explanations on PPT pages
Digital human image	Appearance realism	High, the image is close to real humans, and the details are well-handled
	Style diversity	Diverse, customizable images for different ages, genders, and professional styles
	Naturalness of actions	Natural, smooth movements, and high matching degree with explanation content
Lip synchronization	Accuracy	High, speech and lip movements are basically synchronized, with no significant delay
Voice quality	Voice diversity	Rich, multiple voice tones available to meet different course needs
	Clarity	Clear, no noise or distortion, high voice recognition
Multilingual support	Supported language quantity	Supports common multiple languages such as Chinese, English, Japanese, etc.
	Language conversion effect	Natural conversion, accurate translation, voice and tone match target language habits
Scenarios	Number of scenario types	Rich, including classrooms, laboratories, outdoors, and other teaching scenarios
	Smoothness of scenario switching	Smooth, no stuttering or visual interference during switching
Domestic network	Network access situation	Can be used directly without special network settings
Cost	Development cost	High, customized digital human development costs over 20,000 yuan
	Usage cost	Low, charged based on usage duration or functional modules, good cost-effectiveness
	Maintenance cost	Moderate, regular technical maintenance and content updates are required

3.4. AI-generated digital human process

Typically, creating a digital human involves several steps. Firstly, the target character's persona must be determined, including appearance features, personality traits, and professional identity. Then, technologies such as CV, NLP, and TTS are utilized to process video materials, aiming to achieve the best teaching effect. Next comes the speech synthesis phase, which involves pronunciation training, voice selection, and voice line adjustment. Finally, post-processing steps include dubbing, lip-syncing, and synchronizing sound with the visuals.

3.4.1. Modeling technology

In 1982, American scientist H.C. Perlin introduced the concept of "3D Shape" and applied it to computer graphics. A 3D model refers to the result obtained after rendering, deforming, and stretching a 2D image.

The basic process of 3D modeling includes data acquisition, modeling, texturing, applying textures, and rendering. Among these, the modeling step is primarily completed by modeling software, mainly including 3D modeling tools, texture layers, and more. Depending on different requirements, various modeling methods can be chosen. Currently, popular virtual character production platforms include Unity, Unreal Engine, and Blender, all supporting custom modeling functions. Taking Unity as an example, users can create their favorite character images by dragging and dropping components or importing files. Additionally, pre-made models can be modified. Besides these platforms, some open-source projects also provide free resources and services. For instance, the GPT-3 model launched by OpenAI is a prime example.

3.4.2. Speech synthesis technology

Speech synthesis technology converts human speech signals into digital signals and synthesizes human-like speech through computers. With the advancement of AI technology, TTS (Text-to-Speech) technology has greatly improved. Currently, mainstream TTS tools in the market include Microsoft's TTS, Baidu's Duplex, and iFLYTEK's ASR. Taking iFLYTEK as an example, the company launched an intelligent speech synthesis system based on deep learning technology in 2019, called "iFLYTEK Heard." This system enables speech conversion in multiple languages, including Chinese and English, and supports mixed language recognition. It also features high-quality voice recording, meeting demands in various scenarios. During the production process, pronunciation training for voice actors is conducted first, followed by selecting appropriate voice tones and soundtracks based on teaching content. Lip-sync training and synchronizing sound with visuals are the final steps. In this phase, factors like the network environment and device performance must be considered. For instance, when using a microphone, clarity and noise-free sound are essential, and when playing videos, attention should be paid to syncing visuals with sound. Additionally, issues like intermittent playback or audio-visual inconsistencies should be avoided.

4. Intelligent transformation of educational digital human micro-courses based on AIGC technology

4.1. Development process of traditional micro-courses

The development of traditional micro-courses typically involves multiple stages, including topic selection, teaching design, script writing, material collection, courseware production, recording and explanation, and post-editing. Firstly, teachers determine the theme of the micro-course based on the teaching syllabus and student needs. Then, they carry out teaching design, planning the teaching content and methods. Next, they write a

detailed script, specifying the duration, screen, and lines of each teaching segment. In the material collection stage, teachers need to gather various materials such as pictures, videos, and audio [6]. After that, they produce courseware like PPT and record explanations to present the teaching content in video format. Finally, through post-editing, subtitles, special effects, and other elements are added to enhance the micro-course video.

4.2. Problems of educational digital humans in micro-course production

4.2.1. Technical stability issues

Although AIGC technology and educational digital humans are continuously developing, technical failures such as digital human stuttering, speech recognition errors, and mouth movement synchronization issues may still occur in practical applications, affecting teaching effectiveness. Additionally, the adaptability of teaching scenarios only accounts for 47%, and the average iteration cycle is 35 days per class hour.

4.2.2. Content accuracy issues

AIGC-generated content may contain erroneous or inaccurate information, especially in highly specialized subject areas. When explaining complex knowledge points, educational digital humans may deviate from algorithmic understanding, leading to misleading explanations and confusing students.

4.2.3. Lack of emotional resonance

Despite continuous optimization in expressions and movements, educational digital humans still have a gap compared to real teachers in emotional transmission and resonance with students [7]. Students may find it difficult to feel genuine care and encouragement from digital humans, affecting their learning enthusiasm and initiative.

4.2.4. Homogenization of images

Currently, there is a certain degree of homogenization in the images of educational digital humans in the market, lacking uniqueness and innovation [8]. The appearance, movements, and language styles of digital humans are relatively similar, making it difficult to meet the personalized needs of different schools and courses.

4.3. Transformation path

Firstly, it is necessary to analyze existing educational digital human micro-courses and identify existing problems and deficiencies. For example, many current educational digital human micro-courses still focus on teacher explanations, lacking interactivity and personalization; the content quality varies, making it difficult to meet the needs of different students [7]. Therefore, utilizing AIGC technology through data analysis, intelligent recommendation, and other means can improve the quality and attractiveness of educational digital human micro-courses.

Secondly, it is essential to consider the characteristics of different subjects, types of courses, as well as students' learning habits and needs [9]. Hence, when designing intelligent educational digital human micro-courses, various technical means such as speech recognition technology, face recognition technology, and natural language processing technology can be employed to achieve more intelligent teaching methods.

Lastly, attention should also be paid to the issue of educational fairness, ensuring that every student can enjoy quality educational resources. This requires considering educational fairness when designing educational digital human micro-courses, trying to lower the technical threshold, and enabling more people to participate. Simultaneously, teacher training should be strengthened to better grasp new technologies and improve teaching levels.

5. Conclusion and outlook

In summary, the application of AI technology in the field of education is gradually deepening and has achieved certain results. However, there are still some problems and challenges. For example, there is a significant difference between digital humans and real humans, which may easily cause misunderstanding among students or parents. Additionally, the behavioral capabilities of digital humans are limited, and they cannot completely replace the role of human teachers. Furthermore, digital humans may have a negative impact on students' learning. Therefore, further exploration is needed on how to better integrate AI technology with education, aiming to promote educational equity and improve the quality of education, and drive sustainable development in the education industry. Specifically, research and practice can be carried out from the following aspects: (1) Strengthening research on educational digital human technology and establishing a sound management mechanism and standard specifications; (2) Developing more efficient, convenient, and low-cost digital human production tools to improve production efficiency and quality; (3) Carrying out multidisciplinary cooperation, establishing an interdisciplinary teaching and research team, and developing teaching content and methods suitable for different disciplines; (4) Cultivating interdisciplinary talents with comprehensive literacy and innovation abilities to improve teaching effectiveness and students' comprehensive quality.

Disclosure statement

The authors declare no conflict of interest.

References

[1] Liu B, Nie X, Wang Y, et al., 2025, Generative AI Empowering Education: Technical Framework, Application Fields, and Value. China Educational Technology & Equipment, 2025(3): 61–70.

[2] Zhang Q, Liu R, Xie Y, 2024, Risks and Avoidance Approaches of Generative AI in Higher Education Applications. Journal of Mudanjiang College of Education, 2024(1): 41–45.

[3] Zou T, Xiang L, 2022, AI Applications in the Education and Training Industry: Risk Assessment and Ethical Regulations. Theory and Practice of Education, 42(1): 24–29.

[4] Chen X, Ou Q, 2022, An AI Digital Human Micro-Lecture Production Tool: MetaMaker. China Information Technology Education, 2022(16): 57–58.

[5] Wang A, Liu Z, Liu S, et al., 2024, A Comparative Study on the Application of Multiple Platforms in the Development of Virtual Digital Human Micro-Lecture Resources. Computer Campus, 2024(15): 313–315.

[6] Lin H, 2023, "Virtual Digital Humans" Empower Micro-Lecture Production. Information Technology Education in Primary and Secondary Schools, 2023(4): 78–81.

[7] Wang X, 2023, Research on the Application of Virtual Digital Humans in Micro-Lecture Teaching. Engineering and Technological Research, 8(21): 155–157.

[8] Zhong W, 2024, Research on the Design Strategy of Higher Vocational Japanese Micro-Lectures in the Context of Educational Digitization. Xue Zhou Kan, 2024(25): 114–117.

[9] Zhang X, 2024, Research on AI-Enabled Micro-Lecture Video Production Technology in the Era of Intelligent Education. Scientific Innovation and Application, 14(36): 191–196.

Publisher's note

Bio-Byword Scientific Publishing remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.