Journal of Education and Educational Research
ISSN: 2957-9465 | Vol. 21, No. 1, 2026

Construction and Practice of a Virtual-Real Integrated Practical Teaching Model Driven by AIGC and AI Agents for Higher Vocational Digital Media Education

Zhuohui Zhou *, Cheng Wang

School of Digital Economics, Yibin Industry Polytechnic College, Yibin, Sichuan, 644000, China

*** Corresponding author:** Zhuohui Zhou (Email: 1483296167@qq.com)

Abstract: In digital media programs at higher vocational colleges, students usually learn tools in class but rarely carry a project from a client's brief to a delivered work. Generative AI tools and AI agents have entered professional production chains; in most curricula, though, they remain add-on demonstrations, and the virtual-real equipment in the media center is used mainly for demonstration rather than for graded assignments. This paper reports a redesign of the Photography and Videography course at Yibin Industrial Vocational and Technical College, built around a three-loop model that links online AI-assisted planning, offline multi-device shooting and delivery of industrial tasks. The course ran from March to early July 2026 with 180 first-year students in four classes. Evidence came from a pre/post survey with 176 valid responses, a catalogue of student works, and reviews by Sichuan Tourism Investment Group. Self-rated creative skills rose from 2.3 to 4.1 on a five-point scale, the share of works meeting enterprise acceptance standards rose from 52% to 89%, and 91% of students reported that they produced faster. In total the students completed more than 1,000 photographs and 102 videos. The paper describes the model, the course redesign and the empirical findings, and closes with the conditions under which the model can be transferred to comparable colleges.

Keywords: AIGC; AI Agent; Virtual-reality Integrated Training; Vocational Education; Digital Media; Photography and Videography.

1. Introduction

A promotional video that once needed a small crew is now planned, shot, edited and reviewed by a small team working with generative tools and conventional cameras side by side. Text-to-image, text-to-video, digital human livestreaming and one-click editing have all left the research stage and entered commercial use. What employers now ask for is not fluency in one piece of software but the ability to brief a project, direct an AI-assisted draft, shoot on location and decide when a generated output is ready for delivery. Vocational digital media programs therefore face pressure to teach the whole workflow rather than isolated tools.

Policy has pushed in the same direction. The 2026 "AI+ Education" action plan asks vocational colleges to upgrade traditional majors and build virtual-real integrated learning spaces [1]; the 2026 vocational college AI guidelines go further and specify how generative tools should be placed into the curriculum [2]. At the provincial level, Sichuan targets an AI industry scale of more than 400 billion yuan by 2030 and lists education among its priority application areas [3]. Yibin sits at the receiving end of this push: county-level cultural tourism and agricultural e-commerce projects, including the "Hundreds of Villages, Thousand Broadcasters" livestreaming initiative, keep producing real briefs for short videos, digital human hosts and promotional footage.

What this study addresses is the gap between that demand and the way Photography and Videography is normally taught. In peer programs we observed, generative tools appear as a one-off elective exercise, the virtual-real equipment is reserved for demonstration, and the final assignment is a class film that nobody outside the classroom ever sees. Students graduate familiar with several apps but with almost no practice in taking a project from brief to delivered work. The question we asked was whether the course could be redesigned so that generative tools, AI agents, multi-device shooting and real industrial tasks form one continuous loop instead of separate units.

The remainder of the paper is organized as follows. Section 2 reviews work on AIGC, teaching agents and virtual-real training in vocational education. Section 3 sets out the three-loop model, and Section 4 describes how the course was restructured around it. Section 5 reports the implementation with 180 students and the evidence from surveys, works and enterprise review; Section 6 discusses the findings and their boundaries; Section 7 concludes.

2. Related Work

2.1. AIGC in Vocational Education

Chinese researchers have studied how generative AI reshapes digital resource supply in vocational education and offers personalized practice feedback, while noting that generated content can deviate from the brief and must still be reviewed by humans [4,5]. Broader discussions of ChatGPT and generative AI in this sector point to changes in talent cultivation, teacher roles and evaluation systems, balanced against risks of over-reliance and content authenticity [8,9,20]. Course-level empirical reports are starting to appear. Cao, Hao and Wu, for example, integrated AIGC into an audio-video editing course and measured gains in learning performance [10]; Mi and Jamaludin and Guo have proposed conceptual frameworks for project-based learning with generative AI in Chinese vocational digital media programs [11,12]. Most of this work, however, either stays at the program level or is conceptual; detailed accounts of a photography and videography class that embeds generative tools, with student works and external evaluation, are still rare.

2.2. AI Agents in Teaching

AI agents have grown from simple chatbots into tutoring systems that generate personalized assignments, give real-time feedback and keep a learning record for each student. A 2025 systematic review of 82 studies concluded that hybrid human–AI workflows, in which teachers curate and moderate agent output, outperform fully autonomous tutors [6]. Avatar-based teaching can provide round-the-clock support, but it also raises concerns about over-reliance and the transparency of the teaching relationship [7]. Work such as LearnLM-Tutor has pushed for evaluation-driven refinement and alignment with learning science [19]. In vocational colleges, low-code platforms like Coze allow a single teacher to build a course-specific agent; in practice, however, most such agents stop at question answering and never follow a student through a production workflow from script to final delivery.

2.3. Virtual-Real Integrated Training

Evidence on immersive virtual training has grown quickly. A 2024 systematic review of immersive procedural training in VR reported a significant positive medium effect on learning outcomes, with knowledge transfer showing the largest gains [13]. Randomized studies in organizational training and in bank-clerk vocational preparation found that VR simulations improved conceptual and spatial knowledge, self-efficacy and motivation relative to video-based or conventional classroom delivery [14,15]. Hybrid project-based models for multimedia and animation programs have likewise shown significant gains over conventional instruction [16]. Within Chinese vocational colleges, however, virtual-real practice systems usually treat virtual scenes as pre-built environments; students rarely generate virtual materials themselves and then composite them with their own location footage. The coupling between the generative content pipeline and the virtual-real training studio thus remains underexplored.

2.4. Research Gaps

Three gaps follow from this review. First, few published accounts treat generative tools and AI agents as joint drivers inside one vocational digital media course. Second, the link between virtual-real training and the full generation–production–review–delivery chain is largely missing. Third, empirical work based on real industrial tasks in county-level cities, where first-tier-city infrastructure is unavailable, is almost absent. Reviews of AI education scope in Chinese vocational colleges note that embedded, course-level practice reports remain thin outside computer-science majors [17], and scholarship on generative AI in art and design education calls for pedagogies that keep creative agency in the students' hands [18]. The present study is an attempt to fill these three gaps.

3. Construction of the Teaching Model

3.1. Design Principles

Four design principles guided the model, and they all flow from one premise: students should learn a tool only when a deliverable requires it. Every teaching unit is therefore anchored on a concrete work rather than on isolated software skills. AIGC tools and AI agents act as production partners, but creative direction, aesthetic judgment and ethics review stay with the teacher. Virtual materials generated by AIGC and real footage shot on physical devices are synthesized into a single final work rather than practiced as separate exercises. Finally, works are evaluated against enterprise acceptance standards, and selected pieces are actually delivered to partner enterprises.

3.2. The Three-Loop Model

Figure 1 shows the overall structure: three loops that feed into each other. The online loop rests on AIGC content platforms and teaching AI agents. DeepSeek and Doubao handle brainstorming, script writing and shot-list planning; Jimeng AI and Kling AI generate concept images, style frames and previsualization videos; and a course-specific tutoring agent built on Coze answers questions, reviews scripts, suggests edits and keeps the project file structure organized. In this loop students receive personalized task generation, real-time feedback and a running record of their learning.

The offline loop makes use of the college's media center and a multi-device production pipeline. Students shoot real footage with professional video cameras, camcorders, smartphones and action cameras, then bring that footage back to CapCut (Jianying) and the AIGC tools above to composite it with virtual materials; virtual-real integrated shooting and post-production take place in the VR/AR/MR studio. The tasks for this loop come from Yibin's local cultural tourism and agricultural e-commerce scenarios rather than from invented exercises.

The industrial loop connects the course to real enterprises. Selected works go to enterprise experts for review, are delivered to enterprise accounts for publishing, or are entered into provincial and municipal skills competitions and the "Hundreds of Villages, Thousand Broadcasters" livestreaming project. Evaluation data from this loop flow back to the online loop, so the AI agent can adjust task difficulty and content supply based on what the enterprise actually says.

3.3. Three Curriculum Modules

The curriculum was restructured into three modules, each mapped to one of the loops. Module 1, virtual-real material creation, deals with text-to-image, text-to-video, voice cloning, prompt engineering and style transfer; students generate virtual materials with Jimeng AI, Kling AI and audio generation tools, then judge whether the generated content is usable for real shooting. Module 2, digital human livestreaming design, covers digital human image building, virtual host script design, virtual-real scene integration and livestreaming operation; students write host scripts with Doubao and Coze agents, build digital human avatars, and broadcast from either a virtual studio or a real cultural tourism scene. Module 3, integrated work optimization, covers AI content review, data optimization, work packaging and industrial conversion; students check generated content for false information and compliance, optimize works against enterprise and platform data, and prepare deliverables that meet enterprise acceptance standards.

The three modules follow a four-step progression: AI-assisted ideation and storyboarding, multi-device photography and videography practice, virtual-real material generation and synthesis, and finally project-based work production and evaluation. The ordering is deliberate. No AI tool is introduced on its own; each one appears only when it is needed to advance a concrete deliverable, and the mapping of the three modules to the national vocational job standard is

shown in Figure 2.

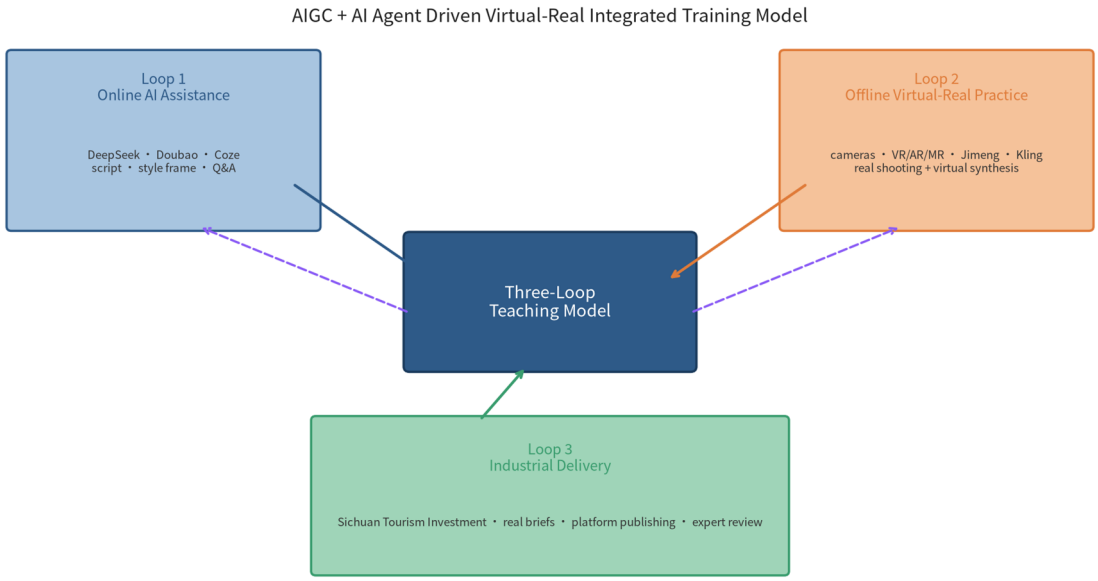


Figure 1. The three-loop virtual-real integrated teaching model

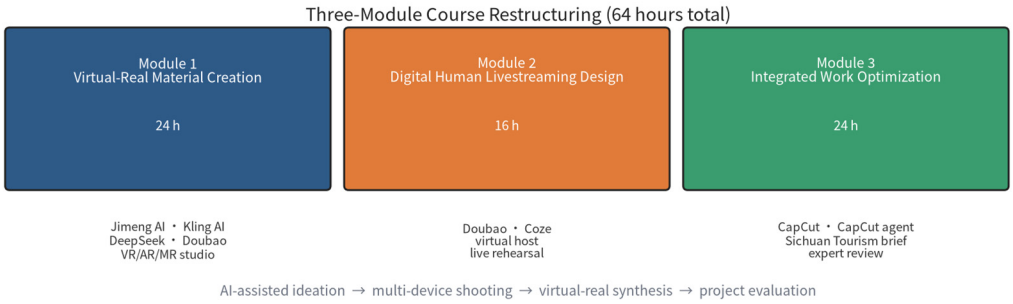


Figure 2. Three curriculum modules mapped to the national vocational job standard

3.4. Evaluation Mechanism

Assessment combines process-based and outcome-based measures. The process part (40%) is recorded automatically by the Coze course agent and through the teacher's observation: task completion logs, the number and quality of AI-assisted drafts, peer review comments and attendance in virtual-real practice sessions. The outcome part (60%) rests on the final work, scored jointly by the teacher for creative and technical quality (30%) and by the enterprise expert for deliverability and market value (30%). Works that pass the enterprise acceptance standard receive a bonus and are actually delivered to the enterprise platform, which gives students a tangible audience and a real portfolio. Figure 3 summarizes the weighting.

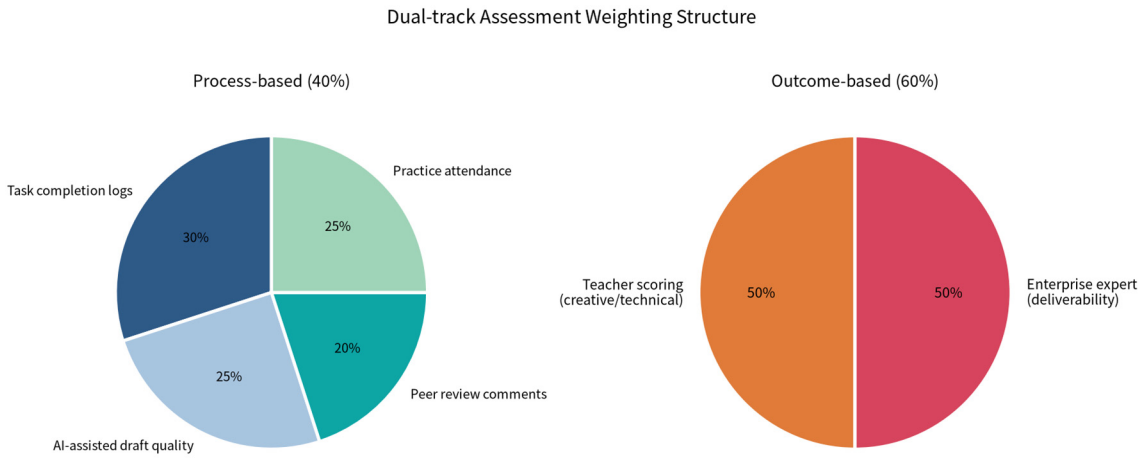


Figure 3. Dual-track assessment weighting structure.

4. Curriculum Restructuring for "Photography and Videography"

4.1. Course Context

Photography and Videography is a core professional course of the Digital Media Technology major, offered in the spring semester of the freshman year. In the 2025 cohort the course ran for four classes totalling 180 students. The reformed version was implemented from March to early July 2026, with 64 instructional hours, half of them practice hours. Figure 4 shows the timeline.

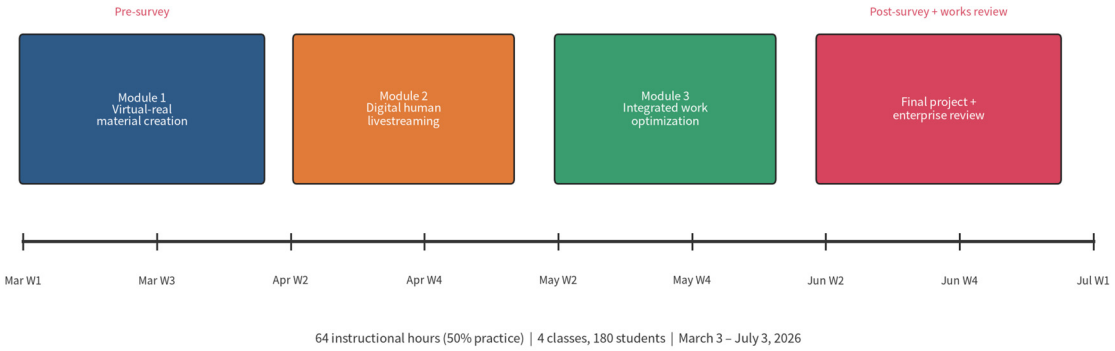


Figure 4. Implementation timeline, March 3 – July 3, 2026

4.2. Tools and Equipment Actually Used

Only tools and equipment actually used in the semester were listed in the syllabus, so that what students practiced matched what the industry uses. On the AIGC and agent side the list is DeepSeek, Doubao, Coze, Jimeng AI, Kling AI and CapCut (Jianying); on the equipment side it is professional video cameras, camcorders, smartphones and action cameras. VR/AR/MR devices in the media center support virtual scene synthesis and digital human livestreaming rehearsal. Table 1 records how frequently students reported using each tool.

Table 1. Tools and equipment actually used, with reported frequency (n = 176).

Tool / platform	Main use in the course	Reported use frequency
DeepSeek	Script and shot-list brainstorming	Every module
Doubao	Host script, dialogue copy, revision feedback	Every module
Coze	Course tutoring agent: Q&A, draft review, file organization	Weekly
Jimeng AI	Concept images, style frames, virtual backgrounds	Modules 1 and 3
Kling AI	Previsualization clips, virtual motion shots	Modules 1 and 3
CapCut (Jianying)	Offline editing, virtual-real compositing, captions	Every practice
Cameras / camcorders / smartphones / action cameras	Real footage on location	Every practice

4.3. Teaching Procedure

Each module runs through six steps, and the same sequence repeats in all three. The industrial partner and the teacher release a real brief, for example a 60-second promotional video for a Yibin scenic spot. Students then move to AI-assisted planning: DeepSeek and Doubao produce the brainstorming, script and shot list. Jimeng AI and Kling AI generate concept images and previsualization. On location, cameras, smartphones and action cameras capture the real footage. In post-production, real footage is composited with virtual materials in CapCut, with editing feedback from the Coze agent. Finally the teacher assesses creative and technical quality while the enterprise expert judges deliverability and market value. Figure 5 illustrates the six steps.

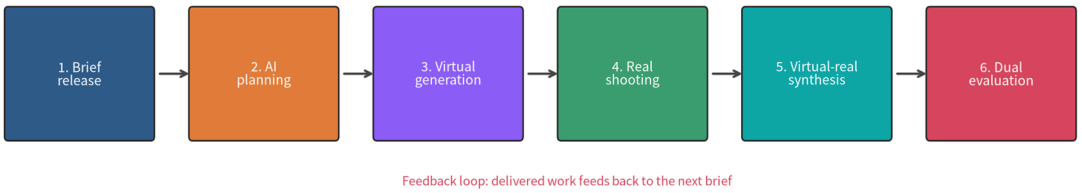


Figure 5. Six-step pedagogical procedure per module

4.4. Data Collection

Four kinds of data were collected. A self-designed student survey went out online at the end of the semester; of 180 copies distributed, 176 valid responses came back (97.8%). The survey covered six dimensions: learning interest, self-rated creative skills, perceived production efficiency, AI tool acceptance, collaboration and overall satisfaction. The same

students also completed pre- and post-course skill self-assessments. All student works were catalogued and categorized. In addition, semi-structured feedback was gathered from Sichuan Tourism Investment Group on the works delivered to its platform.

5. Implementation and Empirical Results

5.1. Student Learning Outcomes

Table 2 and Figure 6 present the main results. On the 5-point self-rating scale, creative skills rose from a pre-course mean of 2.3 to a post-course mean of 4.1. The share of works reaching the enterprise acceptance standard went from 52% in the pilot practice assignment to 89% in the final project. In the end-of-semester survey, 91% of students agreed or strongly agreed that AIGC and AI agents made them more efficient; 87% said the Coze course agent gave timely and useful feedback; 85% wanted to keep using the AI-assisted pipeline in later courses.

Table 2. Pre- and post-course self-rated skill comparison (n = 176, 5-point Likert scale).

Dimension	Pre-course mean	Post-course mean	Improvement
Script and shot-list planning	2.2	4.0	+1.8
Camera operation and composition	2.5	4.2	+1.7
Virtual material generation	1.9	4.1	+2.2
Virtual-real post-production	2.1	3.9	+1.8
Livestreaming operation	2.4	4.1	+1.7
Overall creative confidence	2.3	4.1	+1.8

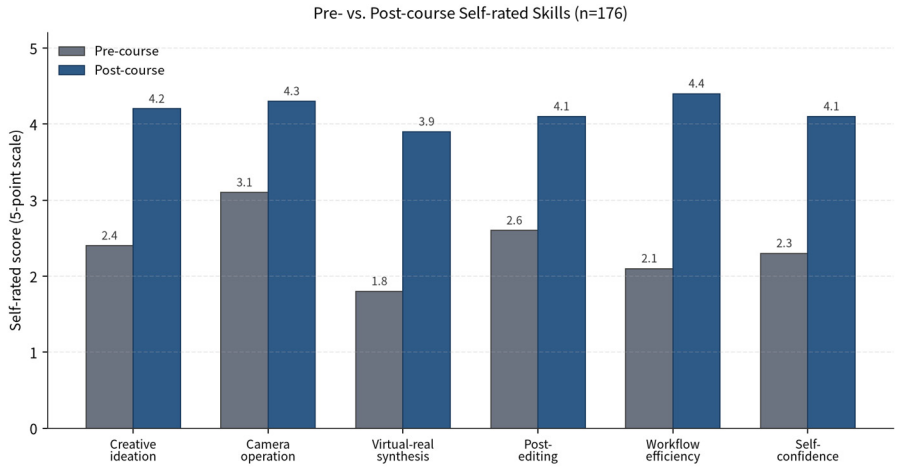


Figure 6. Pre- and post-course self-rated skill comparison

Three survey items deserve separate attention. Asked whether they feared AI would replace their own creative work, 68% of students said they no longer worried by the end of the semester, against 31% at the start; the prevailing view had shifted from "AI will replace me" to "AI will replace those who do not use AI." Asked which part of the workflow benefited most from AI, students ranked script brainstorming (82%), style frame generation (78%) and first-draft editing (74%) highest, while they still felt strongest about camera operation and on-location direction, which they believed AI could not replace. Open-ended comments repeatedly mentioned that the real industrial briefs made the assignments feel meaningful.

5.2. Student Works

Over the semester the students produced more than 1,000 photographs and 102 videos in four categories (Table 3, Figure 7). These were not classroom simulations: most came out of real briefs and real shooting assignments.

Table 3. Student works by category.

Category	Number of works	Typical use case
Photographs	1000+	Campus promotion, event photography, field research, portrait, landscape
Cultural tourism promotional videos	38	Scenic spots and hotels in the region
Campus promotional videos	32	Campus open day and recruitment promotion
Micro-dramas	32	Local cultural stories and public welfare themes

Photography was the most productive category, with more than 1,000 frames across five sub-types: campus promotion, meeting and event coverage, field research, portraits and landscapes. The assignments took students to real locations – school open days, enterprise meetings, field trips to scenic spots and villages, and portrait sessions for their portfolios. Of the 102 videos, 38 were cultural tourism promotional videos produced under briefs from Sichuan Tourism Investment Group and local scenic spots; 32 were campus promotional videos used for the college open day and recruitment; and 32 were micro-dramas built around local cultural stories and public welfare themes.

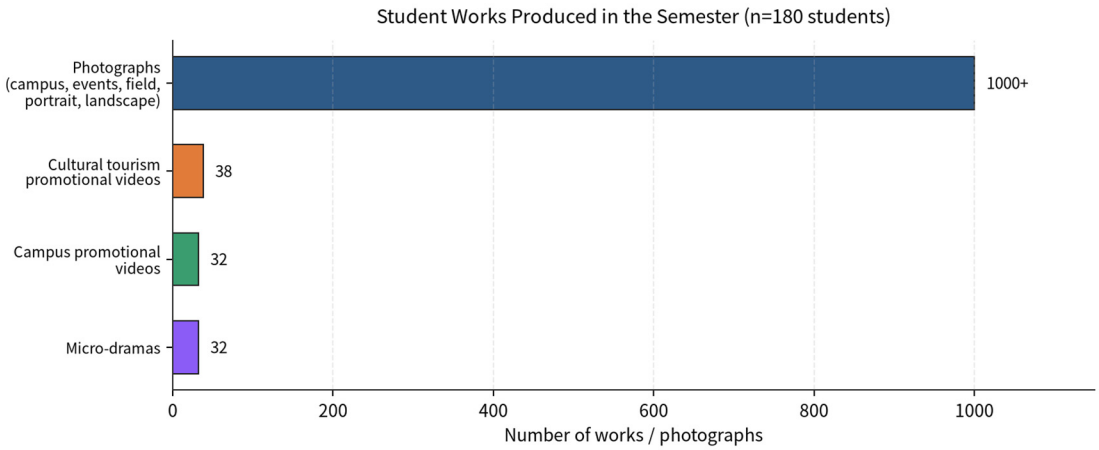


Figure 7. Student works: 1000+ photographs and 102 videos across four categories.

5.3. Enterprise Feedback

Sichuan Tourism Investment Group's content operation team reviewed the delivered works, and their feedback, summarized from semi-structured interviews, falls into three points. On quality, the works were acceptable for platform publishing, and the cultural tourism promotional videos, campus promotional videos and micro-dramas in particular showed clear local characteristics. On efficiency, the human–machine collaborative mode shortened the production cycle noticeably: a 60-second promotional video that used to take a small team two days could be produced by a group of three students in one day, which matters for an enterprise with high-volume content demand. On weaknesses, the team pointed to fine detail – colour consistency across shots, audio level and narrative rhythm – and suggested strengthening these in later iterations.

5.4. Teacher Observation

Two observations stand out from the teacher's side. Participation changed: students who had previously been reluctant to show their work became willing to present, partly because the AI-assisted draft step lowered the fear of a blank page, and partly because work produced for real industrial briefs had a more tangible audience. The teacher's own role also shifted, from demonstrating software buttons to directing creative projects and conducting ethical review; the change was personally demanding but professionally more rewarding.

6. Discussion

The 89% enterprise acceptance rate indicates that coupling generative planning, multi-device shooting and virtual-real synthesis in one production loop is feasible in a vocational setting – not just a technical demonstration. What made it feasible was less the addition of any single tool than the restructuring of the workflow: students did not learn Jimeng AI or Kling AI on their own; they reached for each tool only when a brief demanded a style frame, a previsualization or a virtual host. That is close to how the industry actually uses these tools.

The county-level setting deserves comment, because most published accounts of AIGC in vocational education come from first-tier cities with rich industrial resources. Yibin offered no such infrastructure. What it did offer was local cultural tourism briefs, a county-level e-commerce project, and open-source or low-cost tools. The model appears transferable to similar colleges that can identify one local partner and build on free or cheap platforms rather than investing in proprietary software.

Three limitations qualify these findings. The survey is self-reported and covers one semester; a longer panel and a matched control section would be needed for causal claims. The works were reviewed by a single enterprise partner, and quality varies across students. The agents were general-purpose Doubao and Coze bots rather than models trained on the college's own data, so their feedback on cinematography and local cultural detail was less precise than a teacher's.

A pedagogical tension emerged during implementation. Because the generative draft step lowered the barrier to a first cut, some students submitted works that looked polished but had no coherent narrative. The teacher responded in two ways: every script and shot list had to be approved by the Coze agent and the teacher before any shooting or generation began, and Module 3 added an AI review step in which students had to justify each generated choice with their own creative intent. Introducing AIGC into a production course is therefore not a matter of adding tools; the pedagogical sequence has to be redesigned so that human judgment stays in the lead.

7. Conclusion

This paper reported a redesigned Photography and Videography course built around a three-loop model that couples online AI-assisted planning, offline multi-device virtual-real practice and industrial task delivery, piloted with 180 first-year students over one semester. Self-rated skills, works reaching enterprise acceptance, production efficiency and the catalogue of more than 1,000 photographs and 102 videos all indicate that the model improved both creative confidence and deliverable output. The practical conditions for transfer are a local industrial brief, a low-code agent the teacher can configure, and a workflow in which every AI tool enters only when a deliverable requires it. Subsequent work will extend the model to other courses in the major, fine-tune a course-specific agent on the college's own works, and collect multi-year evaluation data.

Acknowledgments

This work was supported by the 2025 research project of Yibin Industry Polytechnic College (Project No. XJKY2025-03YS). The authors thank the students of the 2025 cohort Digital Media Technology major for their participation.

References

[1] Ministry of Education of the People's Republic of China. (2026). AI+ education action plan. Ministry of Education.

[2] Ministry of Education Vocational College Informatization Teaching Steering Committee. (2026). Vocational college AI application guidelines (2026 edition).

[3] Sichuan Provincial People's Government. (2026). Implementation plan for accelerating the "AI+" pilot program. Sichuan Provincial People's Government.

[4] Song, Y., & Luo, J. (2025). AIGC-empowered digital resource service in vocational education: Vision, challenges and pathways. Chinese Vocational and Technical Education, (7).

[5] Xu, S., & Zhao, T. (2025). Value logic, risks and mitigation of generative AI in vocational talent cultivation. Chinese Vocational and Technical Education, (9).

[6] Córdova-Esparza, D.-M., et al. (2025). AI-powered educational agents: Opportunities, innovations, and ethical challenges. Information, 16(5), 469. https://doi.org/10. 3390/ info16050469.

[7] Fink, M. C., Robinson, S. A., & Ertl, B. (2024). AI-based avatars are changing the way we learn and teach: Benefits and challenges. Frontiers in Education, 9, 1416307. https://doi. org/ 10. 3389/feduc.2024.1416307.

[8] Huang, B., & Jiang, D. (2025). Generative AI embedded in the industry–education integration community: Logic, pathways and directions. Education and Vocation, (12).

[9] Xu, G., Cai, J., Jiang, B., Li, Z., Yang, H., & Zheng, J. (2023). ChatGPT/generative artificial intelligence and the future of vocational education. Journal of East China Normal University (Educational Sciences), 41(7).

[10] Cao, L., Hao, G., & Wu, H. (2024). AIGC enables vocational education teaching: Taking the course of audio and video editing processing as an example. Applied Mathematics and Nonlinear Sciences. https://doi.org/10.2478/amns-2024-3008.

[11] Mi, L., & Jamaludin, K. A. (2025). Generative AI-supported project-based learning in Chinese vocational digital media education: A teacher-regulated conceptual framework. Social Science Research and Review.

[12] Guo, D. (2026). From technical skill to creative agency: Reframing digital media art education in vocational higher education under generative AI. International Journal of Research and Innovation in Social Science, 10(3), 5318–5326.

[13] Jongbloed, J., Chaker, R., & Lavouë, E. (2024). Immersive procedural training in virtual reality: A systematic literature review. Computers & Education, 215, 105038. https://doi.org/ 10. 1016/j. compedu.2024.105038.

[14] Baceviciute, S., López Córdoba, A., Wismer, P., Jensen, T. V., Klausen, M., & Makransky, G. (2023). Investigating the value of immersive virtual reality tools for organizational training: An applied international study in the biotech industry. Journal of Computer Assisted Learning, 39(4). https://doi. org/ 10. 1111/ jcal.12757.

[15] Kickmeier-Rust, M. D., Hann, P., & Leitner, M. (2022). Increasing learning motivation: An empirical study of VR effects on the vocational training of bank clerks. Proceedings of ICEC-JCSG 2022.

[16] Ramadhan, M. S., Jalinus, N., Refdinal, Mulyani, N., & Amin, M. (2024). Development of hybrid project-based learning model for multimedia technology and animation. International Journal of Information and Education Technology, 14(5).

[17] Xie, H., Tao, L., & Xie, L. (2023). Research on prospective scope and approaches of artificial intelligence education in vocational colleges in China. Proceedings of ICELA 2023. Atlantis Press.

[18] Lin, Y. (2025). Curious encounters with AI: A diffractive understanding of art education practices in an era with generative AI. Visual Arts Research, 51(1), 123.

[19] Jurenka, I., Kunesch, M., McKee, K. R., et al. (2024). Towards responsible development of generative AI for education: An evaluation-driven approach. arXiv preprint arXiv:2407.12687. https://doi.org/10.48550/arXiv.2407.12687.

[20] Wu, Q., Li, J., & Xu, G. (2025). Technical characteristics, risks and responses of the talent cultivation evaluation system in vocational colleges under generative AI. Education and Vocation, (3).