A Digital Mentor Apprentice Model for Generative AI-Enabled Practical Training in Rail Transit Smart Operation and Maintenance

Yingge Li*, Hongmei Liu
Guangzhou Railway Polytechnic, Guangzhou, Guangdong, China
**Corresponding Author*

Abstract: Practical training for the intelligent operation and maintenance of rail transit systems has long faced significant challenges, including high equipment costs, difficulty in replicating faults, a shortage of expert mentors, the inability to standardize experience, and a lack of process visibility. Meanwhile, students' use of general-purpose artificial intelligence tools often suffers from issues such as uncontrolled generalization and a disconnect from instructional objectives. To address these dilemmas, this paper proposes and systematically explores the Digital Mentor Apprentice (DMA) educational model. Leveraging generative AI and multi-modal sensing technologies to create a digital mentor, this model provides each student with one-on-one, end-to-end, and traceable mentorship within a training environment that integrates virtual and physical elements, all while adhering to real-world job standards. The model comprises three key components, digital mentor, virtual-physical integrated training environment, and skill benchmarks based on job standards, and operates through a perception-diagnosis-feedback-progression closed loop driven by the dual engines of multi-modal sensing and generative AI. This paper elucidates the model's conceptual essence, core components, operational mechanisms, and implementation pathways, while also designing a representative instructional scenario focused on fault diagnosis for overhead contact line equipment. The DMA model aim to overcome bottlenecks inherent in modern apprenticeship systems, such as the scarcity of mentors and training scenarios and the lack of process visibility, thereby offering a practical implementation path for AI+major curriculum reform in vocational education.

Keywords: Digital Mentor Apprentice Model; Generative AI; Smart Operation and Maintenance in Rail Transit; Practical Teaching Reform

1. Introduction

Driven by the deepening of the new round of technological revolution and industrial transformation, the rail transit industry is rapidly transitioning from a traditional model reliant on manual inspections and empirical judgment to an intelligent operations and maintenance (OM) system based on smart sensing and data-driven diagnostics. Consequently, there is an increasingly urgent market demand for highly skilled, interdisciplinary professionals who possess not only specialized OM knowledge but also the ability to collaborate efficiently with machines [1]. In recent years, the continuous development of intelligent OM systems and key technologies for urban rail transit [2] has placed higher demands on the digital literacy and human-machine collaboration skills of industry professionals. As the primary venues for cultivating skilled talent, higher vocational colleges face a situation where the quality of practical instruction in OM-related majors directly determines how well the talent supply aligns with industry needs. Thus, talent cultivation models focused on intelligent rail transit OM are garnering increasing attention.

Meanwhile, Generative AI, exemplified by tools such as ChatGPT, Doubao, and DeepSeek, has rapidly entered the educational landscape, offering new technological possibilities for addressing long-standing challenges in vocational practical instruction [3]. This trend constitutes a vital component of the digital transformation of vocational education, increasingly driving innovation at both instructional and organizational levels through digital strategies [4]. However, general-purpose large models lack an understanding of industry-specific workflows and job standards, and they do not fully align with educational objectives; if students use them indiscriminately without guidance outside the classroom, these tools can easily become mere means for generating ready-made answers, thereby undermining students' capacity for independent diagnosis and deep thinking. The risk of learner skill degradation due to excessive technological substitution has become a growing concern regarding the application of Generative AI in education [5]. From a broader perspective, researchers at the intersection of AI and vocational education point out that without prudent instructional design, AI may simultaneously give rise to new ones, while solving existing problems [6]. Transforming AI from a mere answer-generating tool into a mentor that guides students through hands-on practical training has become a pressing, real-world challenge in the reform of practical instruction within vocational education. Addressing the practical challenges encountered in specialized practical training for rail transit operations and maintenance, this paper proposes and systematically explores a digital apprenticeship educational model, aiming to provide a feasible and scalable paradigm for the reform of AI + Specialty practical training.

2. Practical Challenges in Smart OM Training for Rail Transit

Apprenticeship is the most representative skill inheritance model in vocational education, which has played a key role in the cultivation of practical ability for a long time. However, rail transit OM equipment is dense, highly dependent on experience, and strict safety requirements. In this field, the traditional apprenticeship system is difficult to meet the current talent training needs. There has been a systematic analysis of the current situation, challenges and improvement direction of virtual imitation training in vocational education, which provides useful references for the innovation of operation and maintenance training models [7].

First, the practical training scene is expensive, dangerous and difficult to reproduce, and it will not be enough for students to practice mobile phones in real life. Core equipment such as contact network, signal, traction power supply, etc. is expensive and consumes energy, and it is difficult to use it for teaching for a long time. Some tasks involve the maintenance of energed, high-altitude or skylight points, which have high safety risks and should not be operated repeatedly by students. Typical faults such as contact network pollution flash and occasional failure of signal equipment mostly occur randomly, and it is difficult to reproduce in accordance with the teaching progress. Therefore, practical training is often mainly based on teachers' demonstrations and students' observation, and students have few opportunities to personally participate in troubleshooting and maintenance.

Second, there is a shortage of high-level practical training teachers, and the traditional apprenticeship system is difficult to scale up. Operation and maintenance skills are highly dependent on practical experience, and many key skills need to be taught on the spot by senior masters or industry experts. However, there is a relative shortage of double-teacher teachers and enterprise experts who have rich engineering experience and are competent in teaching. Dozens of people in a class often have practical training at the same time, and it is difficult for teachers to take into account the progress and differences of each student. Therefore, personalized guidance is difficult to implement, and the teaching results are limited.

Third, a large number of core skills in the field of operation and maintenance belong to implicit knowledge, which is difficult to systematize and teach. Listening to the sound to identify obstacles, looking at the picture positioning, and judging the state of the equipment by the touch all depend on long-term practice and accumulation. It is difficult to fully express such experiences through textbooks, operating procedures or written instructions, and there is a lack of standardized and quantifiable evaluation indicators. Once a senior master or technical backbone leaves his post, his experience is often difficult to preserve and inherit completely.

Fourth, traditional practical training assessments tend to focus on final outcomes while lacking effective documentation and analysis of the entire task-completion process. Current evaluations generally rely on whether students successfully troubleshoot faults, complete repairs, or submit work orders; they lack data regarding the learning process itself, such as fault analysis logic, operational sequences, critical decision-making steps, and the specific stages where errors occurred. Without a record of the entire process, it is difficult for instructors to accurately pinpoint the specific reasons for a student's shortcomings or to devise targeted instructional interventions tailored to individual students.

Fifth, while students widely use general-purpose AI tools, these tools often suffer from uncontrolled generalization and a lack of integration with the actual curriculum. General-purpose large models lack an understanding of rail transit OM, are disconnected from curriculum objectives, and lack traceability in their usage processes; consequently, they risk becoming mere tools for generating answers rather than coaching and guided practice, creating a concern that users might employ AI without actually mastering the skills.

These challenges essentially point to a single core issue: the mentor apprentice model, the most effective approach in vocational education, is scarce in OM practical training and difficult to scale, standardize, or visualize. This presents a clear entry point for reforming the mentor apprentice model through digital technology.

3. Definition and Theoretical Positioning of the Digital Mentor Apprentice Model

3.1 Definition

The digital mentor apprentice (DMA) model proposed in this paper refers to an educational approach in which a digital-mentor, constructed using generative AI and multi-modal sensing technologies, implements a one-on-one, end-to-end, and traceable mentorship process for every student. This takes place within a hybrid virtual-physical OM training environment and is grounded in authentic job standards. It comprises three elements: an AI coaching agent acting as the mentor, a hybrid virtual-physical OM training environment serving as the apprenticeship setting, and rail transit OM job standards acting as the benchmark for craftsmanship. Through the synergy of these three elements, the ancient vocational tradition of mentorship is revitalized via digital technology into a normalized educational mechanism that is accessible to all and leaves a complete digital footprint.

3.2 Distinctions from Traditional and Modern Apprenticeship Models

The traditional master-apprentice system can provide personalized practical guidance, but it is difficult to standardize and replicate and promote. The reason is that tutor resources are scarce, and experience and knowledge are highly implicit. Students' performance often depends on "what kind of mentor they meet". The modern apprenticeship system highlights the key role of enterprise mentors through the dual training framework of school-enterprise cooperation. However, there are still multiple constraints when landing: the number of tutors is insufficient, the cost of on-site practical training is high, the safety risk is high, and it is difficult to provide continuous and personalized guidance for each student under large class teaching [4].

The DMA mode provides a new path to solve these problems. With the help of digital technology, it digitizes the tutor's professional knowledge, practical experience and guidance process, so that it can continue to be used in teaching in a standardized and replicable form. Traditional training relies on limited manual instructors, while DMA breaks through the limitations of time, space and teacher supply. As a result, more students can get stable and consistent learning support, and tutor experience can also be shared on a large scale.

More importantly, DMA not only expands the coverage of tutors. It can also completely record students' learning process and practical performance, and provide richer data for teaching evaluation. Students' operation steps, troubleshooting logic, decision-making process and common mistakes can be collected and analyzed in real time. This promotes the transformation of practical guidance: from qualitative evaluation based on teacher experience to data-based process evaluation and accurate feedback.

In summary, the traditional master-apprentice system solves the problem of "how to teach", the modern apprenticeship system solves the problem of "who teaches", and the DMA model goes further. Its goal is to provide a "digital tutor" for each student, and to make the whole process of guidance recordable, analyzed and evaluated. This not only alleviates the shortage of tutors, but also solves bottlenecks such as limited practical training scenarios and difficulty in quantifying the teaching process, opening up a new way for the cultivation of high-skilled talents in the fields of rail transit operation and maintenance.

4. Core Components and Operational Mechanisms of the DMA Model

4.1 Three-Element Composition

The DMA model is mainly composed of three parts: digital mentors, a practical training environment that integrates virtual and real, and the job specification that integrates the skill standards of craftsmen.

Digital mentors are the core of the whole system. It uses AI intelligence to run through the whole learning process of students, and undertakes the triple functions of demonstration, observation and guidance in different teaching scenarios. At the beginning of the training, it helps students establish correct operational cognition through standardized demonstration. After entering the practice, it continues to track the movements and task completion of the students. Once an abnormality or error is found, targeted prompts and clicks will be given according to the specific situation.

The environment of virtual and real integration provides a platform for the practical teaching of digital tutors. With the help of virtual simulation technology, it can safely reproduce typical faults that are costly, risky or difficult to reproduce in reality. Then combine these virtual scenes with real equipment to form a hybrid practical training environment. Here, students can not only carry out operations close to the real task, but also practice and verify the plan repeatedly under zero risk conditions, so as to steadily improve their practical skills.

Job standards and craftsman skill standards are important cornerstones for digital tutors to carry out teaching and evaluation. The system decomposes job requirements, operation processes and practical experience of senior engineers into observable and evaluable ability indicators, and is matched with corresponding scoring rules and assessment standards. This process transforms implicit experience into explicit knowledge, providing a unified ruler for teaching feedback and ability assessment. It should be noted that this standardization only covers explicit expert experience, that is, rules, procedures, decision-making points and observable behaviors extracted through structured expert interviews and process data analysis. As for those perceptual skills that are difficult to say, such as judging the condition of the equipment by hand feeling, they still need to be cultivated by guiding exercises on physical equipment. In this process, digital tutors play a scaffolding auxiliary role, not a complete replacement.

4.2 Dual-Engine Technical Core: Multi-modal Perception Generative AI

The operation of digital tutors is not only based on generative artificial intelligence, but also on a dual-engine architecture. Multi-modal perception and generative artificial intelligence work together to undertake the two core tasks of environmental understanding and intelligent guidance respectively.

The multi-modal perception engine is mainly responsible for collecting data during the operation of students. It integrates information from fixed cameras, wearable vision devices and various sensors, which can monitor current, voltage, vibration, infrared and other signals. Combined with human posture estimation, target detection, motion recognition and other computer vision technologies, the system can analyze students' actions, tool use, task order and equipment status in real time. On this basis, it judges whether the operation is standardized, identifies key errors, and provides an objective basis for follow-up guidance.

The generative artificial intelligence engine is responsible for teaching interaction and knowledge reasoning. It takes the pre-trained large language model optimized for rail transit OM as the core, and integrates the retrieved enhanced knowledge base, which includes operation standards, typical failure cases and maintenance procedures, so as to provide professional knowledge support for digital mentors. At the same time, the system generates targeted guidance according to specific task situations with the help of prompt templates customized for operation and maintenance scenarios. These guidance include not only operational suggestions, fault analysis strategies and diagnostic analysis, but also heuristic questions. The purpose is to help students establish correct analysis methods, rather than directly giving answers.

It should be emphasized that the diagnostic conclusions and operational suggestions of digital tutors are all based on the retrieved operation standards, standard operation procedures and selected fault case libraries. The generative model is only responsible for organizing, interpreting the retrieval content and asking heuristic questions, and does not make independent judgments. In this way, its output is strictly limited to the controlled knowledge range. When encountering complex or high-risk scenarios beyond the coverage of the knowledge base, the system will start the escalation mechanism and introduce manual mentors and security review to avoid unsafe guidance from model extra-poduction.

Therefore, multi-modal perception provides real and accurate on-site information, and generative artificial intelligence analyzes, reasons and gives teaching feedback accordingly. The two support each other and work together. Compared with the practice of relying only on large language model questions and answers, the architecture can simultaneously understand "what students have done" and "how to guide next", thus improving the intelligence level and teaching effectiveness of digital tutors [8].

4.3 Four-Stage Closed Loop: Perception-Diagnosis-Feedback-Progression

The DMA model consists of four stages: perception, diagnosis, feedback and advancement. The four are interlocked and continuous, forming a continuous closed-loop learning process.

In the perception stage, the system continuously collects the operation data of the students in the mixed reality (MR) training environment. These data cover movement trajectory, tool use, task process and device feedback, which together constitute a detailed record of learning behavior. Then enter the diagnostic stage. The system compares the collected data with the job competence model and standard operating procedures one by one to identify missed inspection, mis-operation, process deviations and weak links, and generate personalized competence portraits and diagnostic results accordingly.

In the feedback stage, the digital tutor gives targeted guidance based on the diagnosis results. Instead of only giving standard answers like the traditional automatic scoring system, it simulates the style of real-life tutors. It guides students to find problems and take the initiative to correct them through questions, demonstrations, error analysis and key concept explanations, so as to improve their ability to solve problems independently.

After this link, the system enters the advanced stage. According to the results of the competence assessment, it dynamically adjusts the training content and the difficulty of the task, recommends the plan that best suits the students' proficiency, and promotes the gradual improvement of skills. With the deepening of training, the system continues to accumulate process data and update the competence portrait in real time to provide an objective basis for formative evaluation and value-added evaluation. As a result, the cycle of practice-diagnosis-improvement-re-practice has been established, teaching feedback and skill development continue to iterate, and the quality of training of rail transit operation and maintenance talents has been steadily improved.

5. Implementation Pathway for the DMA System

In order to promote the implementation of the DMA model in rail transit OM practical training, system development can be divided into five step-by-step stages: establishing a standard framework, developing digital tutors, building a practical training environment, organizing and implementing teaching, and optimizing teaching evaluation. The five stages are interrelated and together constitute a continuously improved operation mechanism.

Establishing Standards First. The first task is to build a unified set of competency standards. The school cooperates with technical experts of rail transit enterprises to sort out the responsibilities, typical work tasks and fault troubleshooting processes of operation and maintenance positions. On this basis, refine the job requirements into observable and evaluable ability indicators, and match them with task evaluation rules and scoring criteria. As a result, there is a unified basis for digital tutor development, practical training design and learning evaluation. At this stage, it is indeed necessary to invest scarce senior expert resources, but it is a one-time preliminary investment. Once expert experience is transformed into a standardized guide, it can be reused and shared on a large scale among different batches and different campuses. The role of experts has also changed: it is no longer to continuously guide individual students, but to regularly formulate and verify standards.

Deploy digital mentors. Based on the ability standard, build a digital mentor system specially customized for rail transit operation and maintenance. With artificial intelligence (AI) intelligence as the core, it integrates the field knowledge base, prompt template, multi-modal identification module and teaching strategy to realize functions such as professional knowledge retrieval, teaching interaction and learning analysis. At the same time, the digital tutor is deployed together with the school's teaching platform and practical training system to form an intelligent teaching support platform that can be flexibly configured and updated sustainably.

Build a practical training environment. Relying on existing laboratory equipment and digital teaching resources, we can create a practical training environment where virtual simulation and physical equipment operate together. Design a comprehensive practical training process around typical operation and maintenance tasks, and establish a case resource library covering a variety of failure types and difficulty levels to provide students with practical conditions close to real work scenarios. This hybrid mode not only ensures the authenticity of practical training, but also reduces equipment loss and safety risks as much as possible, so that students can practice repeatedly and explore independently.

Closed-loop practical training. In practical training, teachers, digital tutors and students work together. Teachers are mainly responsible for organizing teaching, supervising the process and controlling key links; digital tutors provide dynamic guidance and feedback based on students' real-time operations. Students carry out troubleshooting, equipment maintenance and operation training around specific OM tasks, and promote the process of perception-diagnosis-feedback-advance to gradually improve their job practice skills.

Data-driven evaluation. The digital platform automatically records the learning behavior of students in the whole process of practical training, collects data such as operation process, task completion, prompt use, troubleshooting results and task efficiency, and carries out formative evaluation and value-added evaluation according to the job competence standards. The evaluation results are not only used to analyze students' ability development, but also to provide a basis for task design adjustment, digital tutor knowledge update and teaching strategy optimization. As a result, the system can be continuously iterated to form a closed loop integrating teaching, evaluation and platform optimization.

6. Design of a Typical Teaching Scenario: Practical Training on Catenary Equipment Fault Diagnosis

This article selects the fault diagnosis and processing tasks of contact network equipment as a typical case. Based on the closed-loop mechanism of perception-diagnosis-feedback-advance of DMA model, a complete teaching implementation process has been designed.

In the perception stage, teachers issue fault tasks on the platform according to learning goals, such as abnormal power supply in a certain contact network segment. Students enter the practical training environment of virtual and real integration, and carry out equipment inspection, status analysis and fault troubleshooting according to the real operation and maintenance process. At the same time, digital tutors use multi-modal perception technology to collect multi-source data, including inspection paths, operation steps, tool use and device status changes, so as to generate dynamic records of students' operation behavior.

Entering the diagnosis stage, the digital tutor analyzes the collected process data. It compares the data with the job competency model and the standard operating procedure (SOP) item by item to judge whether the operation is standardized. The system can accurately locate the links of the problem, such as missed inspection, misjudgment or process errors, and generate corresponding competency analysis results. For example, if students are eager to determine the cause of the failure before completing the key equipment inspection, the digital tutor will identify the defects in their diagnosis logic and provide a basis for follow-up guidance.

In the feedback stage, the digital tutor simulates the teaching style of senior operation and maintenance experts and gives targeted guidance based on real-time performance. It does not give the correct answer directly like traditional training, but focuses on heuristic guidance, prompting students to re-examine the fault phenomenon and analysis logic through a series of follow-up questions. For example, it will ask: "What is the basis for you to determine the abnormal power supply?" Have factors such as insulator staining and flashing been excluded?" In this way, students can find problems independently and optimize diagnosis and reasoning.

As the training deepens, the system enters the advanced stage. Digital tutors dynamically adjust the difficulty of the task according to the training data and proficiency accumulated by students, and guide them to gradually transition from simple failures to complex failures, multi-factor abnormalities and even high-risk operations. For example, after mastering the basic troubleshooting method of power supply abnormalities, students can advance to comprehensive scenarios involving compound faults or live operation, so as to improve the problem-solving ability under complex working conditions. During the practical training, the operation behavior, diagnosis process, interaction record and task completion details are all recorded by the platform in real time and used to build students' professional competence portraits. On the one hand, these data are convenient for teachers to track skill development and carry out personalized teaching, and on the other hand, they also provide support for formative evaluation and value-added evaluation.

This case shows that the DMA model is not just a theoretical framework for digital teaching. It can be combined with specific operation and maintenance tasks and integrated into the actual teaching process. With the help of the perception-diagnosis-feedback-advance workflow, digital tutors provide continuous support for students' practical training, and also open up a practical digital teaching path for the cultivation of high-skilled talents in the field of rail transit.

7. Expected Outcomes and Value for Scaling

The application value of this model is mainly reflected in three aspects. First, it alleviates the shortage of resources for traditional apprenticeship tutors through digital tutors providing continuous and personalized guidance. Second, it transforms expert experience into reusable standardized teaching resources with the help of job competence model and task evaluation standards, thus promoting the transmission of professional knowledge [9]. Third, it relies on the process data of perception, diagnosis, feedback and advancement stages to dynamically evaluate students' skill development, thus supporting the optimization of accurate teaching and talent training mode. Related research has proposed a closed-loop framework for practical computer vision training in rail transit operations and maintenance [10], providing a theoretical reference for the mechanism design of this model. This model prioritizes a lightweight implementation approach, focusing on encapsulating existing large model APIs and reusing established training and platform resources. It offers excellent adaptability for scaling and is expected to be transferable to other technical and vocational disciplines-ranging from rail transit operations and maintenance, such as power supply, signaling, and rolling stock systems to a broader spectrum of technical fields-thereby providing a practical framework for AI+major curriculum reform in vocational education.

8. Conclusion

Driven by the dual forces of the smart OM transformation in rail transit and the rise of generative AI, the DMA model proposed in this paper deeply integrates the traditional vocational mentor apprentice training method with digital technologies. Centered on three key elements-digital mentors, a virtual-physical integrated training environment, and job-standard-based skill benchmarks-and a four-stage closed-loop process, perception, diagnosis, feedback, and progression, this model explores new pathways to overcome the bottlenecks of the modern apprenticeship system, specifically the scarcity of mentors and training scenarios, as well as the lack of process visibility. Future research will involve continuous validation and iteration through broader teaching practices, further refining competency maps, evaluation metrics, and scaling mechanisms to bring the model to maturity.

Acknowledgments

This work was supported by the Guangzhou Railway Vocational and Technical College Newly Introduced Talents Scientific Research Start-up Project (No. GTXYRC250105).

References

[1] Ma Jun, Ma Junqi. Connotation, constituent elements, and cultivation strategies of digital skills for vocational college students from the perspective of digital labor. Journal of Southwest University (Social Sciences Edition), 2024,

50(1):224-236.
[2] Cai Yujing, Gao Fan, Meng Yukun, et al. Design and key technology research of intelligent operation and maintenance systems for urban rail transit equipment. Railway Computer Application, 2023, 32(7):79-83.
[3] Gu Xiaoqing, Hu Yiling, Hao Xiangjun. Re-examining artificial intelligence and the future development of education amidst the ChatGPT craze. Academic Digest of Liberal Arts in Higher Education Institutions, 2023, 40(5):152-153.
[4] Han Xibin, Li Mixue, Guo Wenxin. New breakthroughs in empowering vocational education through digital strategies: New progress in research and practice of vocational education digitalization in 2024. China Vocational and Technical Education, 2025(2):39-48.
[5] Wang Youmei, Wang Dan, Wang Haijie, et al. The deskilling crisis and countermeasures regarding the educational application of generative AI: An analysis based on the framework of transformative reverse thinking. Open Education Research, 2025, 31(4):97-108.
[6] An Hong, Lu Fali. Artificial intelligence and vocational education reform: Research review and potential issues. Teaching in Forest Regions, 2023(10):37-42.
[7] Han Xibin, Shi Wanruo. Analysis of dilemmas and resolution strategies for virtual simulation training in vocational education. e-Education Research, 2025(9).
[8] Lu Yu, Tang Xiaoyu. Morphological hierarchies and progression paths of classroom teaching empowered by generative AI. e-Education Research, 2025,46(6).
[9] Liu Bangqi, Nie Xiaolin, Wang Yafei, Yuan Tingting, Zhao Ziqi, and Zhang Guoqiang. Generative AI Empowering Education: Technical Framework, Application Scenarios, and Value. China Educational Technology, 2025(3).
[10] Li Yingge, Guo Runhao, Liu Hongmei, et al. A computer vision-enabled closed-loop framework for adaptive practical-skill instruction in higher vocational education: a rail transit maintenance case. Occupation and Professional Education, 2026, 3(6).