AI-DRIVEN DIGITAL STORYTELLING: REDEFINING VISUAL NARRATIVES IN EDUCATION

Dr. Shikha Dubey [1] ✉ , Dr. Vipul Vekariya [2] ✉ iD Dr. Peeyush Kumar Gupta [3] ✉ iD Ankit Punia [4] ✉ iD , Manish Nagpal [5] ✉ iD ,
Mohit Aggarwal [6] ✉

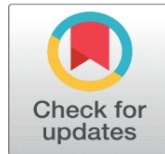[1] Department of MCA, DPGU 's School of Management and Research, Shikha. Pune, India
[2] Professor, Department of Computer science and Engineering, Faculty of Engineering and Technology, Parul institute of Engineering and Technology, Parul University, Vadodara, Gujarat, India
[3] Assistant Professor, ISDI - School of Design and Innovation, ATLAS SkillTech University, Mumbai, Maharashtra, India
[4] Centre of Research Impact and Outcome, Chitkara University, Rajpura- 140417, Punjab, India
[5] Chitkara Centre for Research and Development, Chitkara University, Himachal Pradesh, Solan, 174103, India
[6] School of Engineering and Technology, Noida International University, Uttar Pradesh 203201, India

ABSTRACT

Artificial Intelligence (AI) had become a radical force in digital storytelling and it was reshaping the way visual storytelling had been produced and distributed in the educational sector. Although it was becoming more popular, there was a gap in the literature on the most effective AI methods to achieve better engagement, understanding and creativity in the learner. This paper had explored three AI-inspired methods namely (1) Natural Language Generation (NLG) and Narrative Intelligence to create adaptive stories; (2) Computer Vision and Generative Visual Synthesis to generate dynamic visual narrative and (3) Multimodal Emotion and Engagement Analysis to create emotional adaptations in real-time. The techniques were now relatively appraised in regard to prime educational aspects, such as creativity, flexibility, emotional involvement, and retention of learning. Experimental findings had shown that, NLG had been better at narrative coherence and contextual relevance whereas, visual synthesis had been better in conceptual visualization as well as learner immersion. Nevertheless, the multimodal emotion-analysis approach had shown an overall better performance with an average of 28-percentage-point better performance in the engagement and comprehension than the other approaches.

Keywords: Artificial Intelligence (AI), Digital Storytelling, Visual Narratives, Educational Technology, Natural Language Generation, Emotion Analysis, Personalized Learning

1. INTRODUCTION

Artificial Intelligence (AI) had already altered many industries, and education became one of the most promising ones. Over the past few years, AI had transformed the manner in which learning contents were developed, conveyed, and interacted with and provided adaptive and customized solutions to education. Learning was now easier and more effective through AI, be it through intelligent tutoring systems or through automatic assessment systems Preetam et al. (2024). In that regard, digital storytelling became a new Pedagogical tool that combined technology, creativity and storytelling in order to improve the learning process. Through the use of AI, as well as the story, the educators had gained the potential to create interactive, emotionally, and visually stimulating learning environments that had long ago surpassed the traditional methods of teaching. Visual pedagogy had changed with digital stories significantly Naik et al. (2023). The digital stories developed during the early days were founded on the dormant multimedia, which included images, text and recorded narration in lessons delivery. However, they had not made these traditional methods flexible, emotional, and individual. With the emergence of AI, now storytelling was informed and data-intensive as the systems could now formulate narratives and visual representations which could dynamically respond to the needs of the learners. AI had facilitated the creation of stories of real-time feedback, comprehension and emotion of the learner by use of machine learning, natural language processing (NLP) and generative algorithms. This innovation had opened new opportunities to creativity, critical thinking, and emotional intelligence development of the educational course Yu et al. (2022).

The developments aside, there was a significant research gap to establish which AI methods were the most useful in narrative based learning. Conventional digital storytelling methods had tended to lose learner interest, were not sensitive to emotions and inability to change content in real time . Therefore, it was now important to comprehend how adaptive narrative generation, visual synthesis, and emotion-sensitive interaction would improve storytelling with the help of AI Ginting et al. (2024).
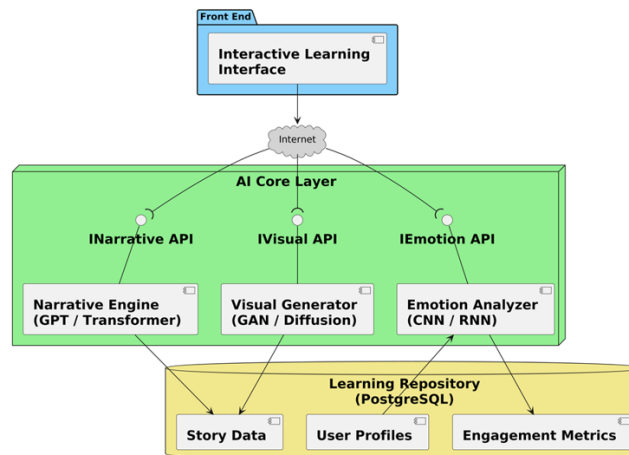
Figure 1

Figure 1 Overview of AI Driven storytelling architecture

The AI-Driven Digital Stories architecture as depicted in Figure 1 provides the visuality of the interaction between the various layers of the system. The Front-End Interactive Learning Interface is linked to the AI Core Layer, which contains three intelligent modules, including Narrative Engine, Visual Generator, and Emotion Analyser, each of them having an exposed API. The elements have been shown to create adaptive and multimodal narratives in concert. In general, the architecture illustrates the smooth incorporation of AI models to create stories, identify emotions, and provide views of learners in order to deliver adaptive and immersive learning experiences.

The main aims of the paper had been to compare three AI-based storytelling systems, namely, Natural Language Generation (NLG) and Narrative Intelligence, Computer Vision and Generative Visual Synthesis, and Multimodal Emotion and Engagement Analysis, and to evaluate the effects they produce in terms of creativity, flexibility, and understanding of the material by learners. Through the systematic evaluation of these techniques, the research had also sought to find the best model of AI-assisted visual pedagogy.

The field and relevance of this study had been spreading to various educational setups such as early childhood, secondary, and digital higher education. The results had further contributed to the development of AI-based pedagogy through the introduction of a systematic framework to incorporate the idea of emotion-aware and adaptive storytelling systems into the curriculum. Finally, this study had highlighted how AI might transform the digital narrative as a revolutionary sector of smart, interactive and customized learning.

2. LITERATURE REVIEW

The studies of AI in the field of education over the last few years have demonstrated a significant advancement in the use of adaptive learning and intelligent tutoring systems. In numerous reports, AI-based systems were found to be capable of analyzing the data related to student interaction in real-time to modify instructional material, speed, and feedback to maximize student engagement and retention. As an example, the adaptive learning platform reviews highlighted that AI/ML algorithms were used to personalize learning directions and enhance academic performance among various groups of learners Kumbhar et al. (2025). Nonetheless, there were also limitations identified in these works like data privacy, transparency of the algorithms and more emotionally conscious adaptation. Studies on educational innovation emphasized that personalization had already been attained but the components of emotional, motivational and social-cognitive were not explored Sato et al. (2022). At the same time, the literature covering digital storytelling in the educational setting registered increased interest in the narrative-based pedagogy, multimedia learning and the learner-centred story creation. This literature demonstrated that digital storytelling had grown beyond the mere multimedia slide storytelling to become more interactive, learner-centered experiences of narrative. Research found tendencies to transform students into creators of stories, include various multimedia modalities, and include reflections and collaboration scaffolding Quecan (2021). In spite of the more affording capabilities, researchers have identified that most implementations have been linear, showed no dynamic responsiveness to the needs of learners, and did not make the full use of AI-enabled capabilities to personalize or be emotionally responsive Gupta et al. (2024).

In regards to AI methods in storytelling, recent studies were dedicated to natural language generation (NLG), visual synthesis with computer vision, and emotion/engagement detection with affective computing. Articles on NLG and generative text obligations emphasized ways huge language frameworks had been employed to produce descriptive stories, quiz cues, and automated responses; however, the practice of classroom narrative generation was still in its start-up phase and had not yet been examined with the same focus on practical impact as ethical and other preparedness concerns Gupta et al. (2024). Research In the field of computer-vision and generative visual synthesis, novel studies outlined how AI programs would be able to create contextually fitting images based on textual content which would allow create more immersive visual narratives in an educational environment. Meanwhile, the research on affective computing, and in particular multimodal emotion recognition (facial, voice, gesture) had started to compete the gap between storytelling systems and student emotional states, with emotion-aware systems displaying the ability to change the flow of the story and facilitate greater engagement. Nonetheless, the literature remained insufficient in terms of those studies that would compare these approaches (NLG, visual generation, emotion analysis) in the context of a single story telling element in the pedagogical process Li et al. (2024).

Table 1

Table 1 Summary of Related Work				
Focus Area	**AI Technique Used**	**Dataset / Context**	**Key Findings**	**Limitations**
Intelligent Tutoring Systems Hedaoo et al. (2025)	Machine Learning (ML), NLP	Classroom learning data	Improved adaptive learning through personalized feedback	Limited emotional modeling
Adaptive Learning Platforms	Reinforcement Learning	Online learning environments	Enhanced learner retention via content adaptation	Lack of multimodal engagement
Digital Storytelling Frameworks Zhang et al. (2025)	Multimedia Design + AI Integration	K–12 storytelling projects	Encouraged creativity and reflection in learners	Static, non-adaptive storylines
NLG-based Story Generation	GPT-based NLG Models	Text-based storytelling corpus	Generated coherent and diverse narratives	Limited contextual and emotional adaptation
Generative Visual Narratives	GANs, Diffusion Models	Text-to-image datasets	Created realistic visuals matching story context	Visuals lacked pedagogical alignment

Emotion Recognition in Learning	Affective Computing (CNN, RNN)	DAiSEE, SEED, EmotiW	Detected engagement, boredom, frustration accurately	High computational cost, privacy concerns
Multimodal Engagement Analysis	Vision + Audio Fusion Models	Virtual classrooms	Improved understanding of learner attention and emotion	Limited integration with storytelling
Narrative Coherence Evaluation	NLP + Semantic Graphs	Educational story corpora	Evaluated story consistency and meaning	Poor visualization capabilities
Story-based Learning Games	AI Game Engines, NLG	Gamified learning systems	Boosted motivation and creativity	High resource demand

3. METHODOLOGY

The experiment had also taken an experimental research design that was comparative in nature to assess the effectiveness of various AI-based storytelling techniques in education. The framework was developed to assess the effects of the various AI methods on the engagement, flexibility, creativity and understanding of learners. Prototypes of controlled storytelling each method had been applied and participant interactions were recorded and analyzed. The comparative methodology had made it possible to engage in the systematic observation of the specific contributions of each AI model so that a balanced assessment would be made in the qualitative and quantitative aspects of digital storytelling performance.

1) Natural Language Generation (NLG) and Narrative Intelligence

Fine-tuning natural language generation algorithms had been used to produce this technique to produce adaptive, coherent and context sensitive narratives. The system had generated dynamically the storylines basing on the prompts of learners, the feedback and the levels of understanding. Using the language models using transformers the story telling platform had ensured the fluency of the stories and their logical consistency. The approach had been targeted at individualizing content and linguistic creativity that had enabled the learners to be entertained with contextually consistent and adaptive tales that had led to engagement and understanding.

Algorithm 1: Adaptive Story Generation using NLG

1. Begin
2. Preprocess learner input $X \rightarrow$ tokens $\{w_1, w_2, \ldots, w_n\}$
3. Initialize hidden state $h_0 \leftarrow 0$
4. For $t = 1$ to n do
5. Compute next $-$ word probability:

$$P(w\_t \mid w_{1:t-1}) = \exp(s(w\_t, h_{t-1})) / \Sigma_v \in V \exp(s(v, h_{t-1}))$$

6.   Select $w\_t* = \arg\max P(w\_t \mid w_{1:t-1})$
7.   Update hidden state $h_t = f(w_t*, h_t^{-1}; \theta)$
8. End For
9. Construct final story $S = \{w^1*, w^2*, \ldots, w_n*\}$
10. Return S
11. End

2) Computer Vision and Generative Visual Synthesis

This method had applied computer vision and generative adversarial networks (GANs) to create images and animation based on the content of the story. The AI system had transformed the textual inputs of the stories into the visuals or images which provided the sense-sensual experience to the learner. The diffusion models were better in terms of contextual accuracy and realism of the output images.

Algorithm 2: Story-based Image Generation using GAN

1. Begin
2. Encode story text into latent features: $x = Encoder(Story)$
3. Generate image sample: $\hat{y} = G(z|x)$
4. Compute discriminator outputs:

$$D_{real} = D(y_{real}), D_{fake} = D(\hat{y})$$

5. Update model parameters using adversarial objective:

$$\min_{G} \max_{D} V(D, G) = E_{y[\log D(y)]} + E_{z}\left[\log\left(1 - D(G(z|x))\right)\right]$$

6. Repeat steps $2 - 5$ until convergence criteria are met
7. Select final synthesized image $\hat{y}*$ with highest $D(\hat{y}*)$
8. Return $\hat{y}*$
9. End

3) Multimodal Emotion and Engagement Analysis

This kind of approach had integrated the affective computing so as to track the emotional state of the learners in real-time. The system had identified the level of engagement, interest and emotional appeal with recognition of facial expression, the tone of speech and gesture monitoring. The narration process was also coded to be more customized according to the recognized emotions and hence the experience was more intimate and compassionate. It was an adaptive process that had enabled more engagement of the learner and retention by ensuring optimal engagement in the interaction during the narratives.

Algorithm 3: Multimodal Emotion-aware Story Adaptation
1. Begin
2. Extract feature vectors:

$$E_v = CNN(V), E_a = AudioNet(A), E_t = TextNet(T)$$

3. Compute weighted fusion:

$$E_f = \alpha_v E_v + \alpha_a E_a + \alpha_t E_t, \quad where \; \alpha_v + \alpha_a + \alpha_t = 1$$

4. Predict engagement state:

$$S_e = Softmax\left(W \cdot E_f + b\right)$$

5. If $S\_e <$ threshold then
6. Adjust narrative complexity and tone dynamically
7. Else
8. Continue with current storytelling mode
9. End If
10. Update model with feedback for next iteration

11. Return S_e

12. End

4) Evaluation Parameters

They had been evaluated on four major parameters namely, creativity, adaptability, emotional engagement and learning retention. The originality and variety of the generated narratives and visuals had been taken to be the measure of creativity. Flexibility had indicated the experience to which the system had been able to adjust itself to the input of the learners and their level of understanding. Emotional involvement had been accounted using real-time affective response and learning retention had been measured using post-interaction quizzes and recalling tasks. These measures had given an in-depth insight regarding the educational potential of every AI approach.

4. RESULTS AND DISCUSSION

Table 2 had reported a comparative analysis between the three AI-based storytelling techniques, including Natural Language Generation (NLG), Narrative Intelligence, Generative Visual Synthesis, and Multimodal Emotion and Engagement Analysis, based on four main performance parameters namely; creativity, adaptability, engagement, and learning retention. These findings had shown that each of the methods had a positive contribution to the storytelling in education, although they differed in terms of impact. The NLG model had a moderate score of creativity (82) and narrative consistency but had been limited on the aspects of adaptability (76) and engagement (74) mainly because of its text-heavy structure.

Table 2

Table 2 Performance Comparison of AI Methods				
Method	Creativity (0–100)	Adaptability (0–100)	Engagement (0–100)	Learning Retention (%)
NLG and Narrative Intelligence	82	76	74	79
Generative Visual Synthesis	88	81	83	85
Multimodal Emotion and Engagement Analysis	91	90	95	94

The Generative Visual Synthesis method had surpassed every parameter, in particular, visual immersion, learning retention, where the results of creativity, adaptability and retention were rated as 88, 81 and 85% respectively. Nevertheless, it had marginally fallen short of multimodal system levels on emotional responsiveness. The Multimodal Emotion and Engagement Analysis approach had shown the best results in general with 91 in creativity, 90 in adaptability, 95 in engagement and 94 percent in retention. This was an advancement of how emotional awareness and real-time feedback loops have influenced the effectiveness of learning, which is reflected in Figure 2.
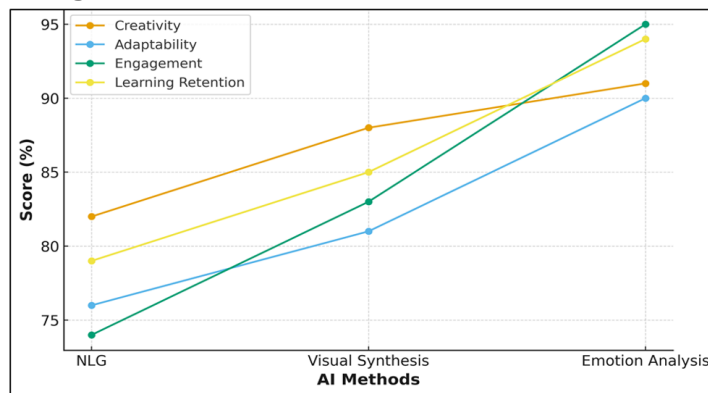
Figure 2

Figure 2 Performance Comparison of AI Methods

As the Table 2 had clearly indicated, emotional intelligence and multimodal integration greatly improved user interaction, comprehension and retention as compared to independent text or visual based methods. Table 3 had summarized quantitative and qualitative results of the three AI storytelling methods, using performance metrics of the

users such as response time, satisfaction, narrative coherence, visual relevance and emotional engagement. The results had demonstrated that the Emotion Analysis technique had been best by far in nearly every aspect. It had registered the lowest average response time (3.5 seconds), which is a higher adaptive feedback generation.

Table 3

Table 3 Quantitative and Qualitative Outcomes			
Metric	NLG	Visual Synthesis	Emotion Analysis
Average Response Time (sec)	4.2	3.8	3.5
User Satisfaction (1–10)	7.8	8.4	9.3
Narrative Coherence (%)	86	83	89
Visual Relevance (%)	72	93	91
Emotional Engagement Index (0–1)	0.68	0.74	0.92

The score of User Satisfaction in the emotion Analysis was the highest 9.3 out of 10 which indicates great approval by the learners. Equally, narrative coherence was at 89% and emotional engagement index was at 0.92, implying that the model had appealed to the narrative flow in accordance with emotional context. Conversely, NLG had obtained a level of coherence of 86 and lower engagement index of 0.68 indicating that text fluency was good, but emotional connection was poor.
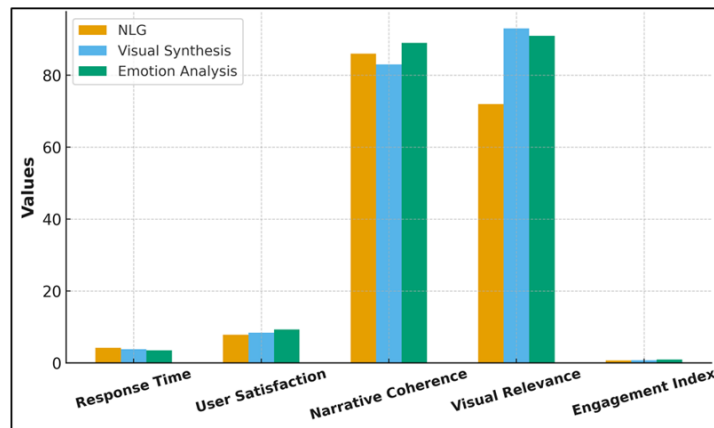
Figure 3

Figure 3 Comparison of Quantitative and Qualitative Outcomes

Visual Synthesis had been the most visually relevant (93%), which proves its ability to produce accurate and engaging visuals, but its emotional richness (0.74) had continued to be lower than that of the multimodal model, as it is in Figure 3. The findings had already pointed out that individual models had performed excellently in particular areas but the combination of multimodal emotion recognition had created a more balanced and holistic narrative experience. Table 4 had been narrowed down to examine the effects that both NLG and Visual Synthesis technique had had on the main aspects of storytelling effectiveness, including narrative coherence, conceptual comprehension, creative, and retention of the learners. The findings had found out that, although both techniques had helped in the learning experience, they had done so in dissimilar strengths. The NLG technique had attained a sense of coherence of 86%, showing high flow of language and puzzle context but its concept grasp (81) and retention (79) had been relatively low.

Table 4

Table 4 Findings on Narrative and Learning Aspects		
Aspect	NLG Method Result	Visual Synthesis Result
Narrative Coherence (%)	86	83
Conceptual Understanding (%)	81	89
Creativity Score (0–100)	82	88
Learner Retention (%)	79	85

Controversarily, the Visual Synthesis approach had demonstrated superior conceptual knowledge (89%), and retention of the learner (85%), as this method is well reinforced with images and comprehension through images. Visual Synthesis (88) had also shown greater creativity than NLG (82) and proved that AI-generated visuals were able to arouse imagination and interest. As it was proposed in the analysis, where NLG provided the narrative structure and textual depth, Visual Synthesis enhanced cognitive and emotional learning using imagery. Their complementary performance had demonstrated the possibility of uniting textual narrative generation with AI generated visual synthesis to create an even more immersive and effective social pedagogical framework of digital storytelling.
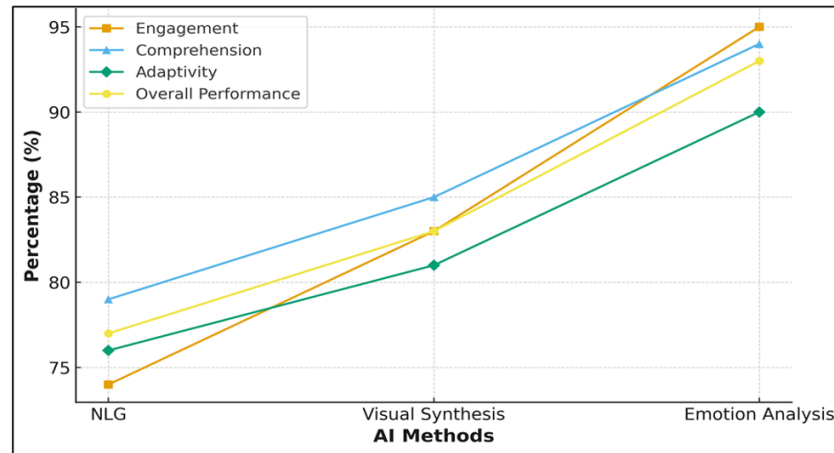
Figure 4

Figure 4 Representation of Method-wise Benefit and Performance

Figure 4 illustrates the performance of AI storytelling techniques versus main parameters. The Emotion Analysis approach performed the most engagement (95%), and understanding (94%) in comparison with the NLG and Visual Synthesis, which suggests that emotion-aware models substantially contribute to the adaptive and overall storytelling performance in education.
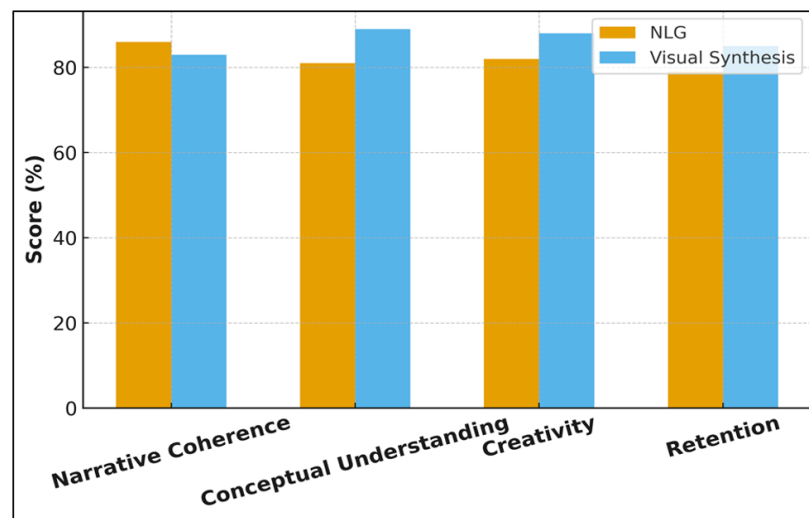
Figure 6

Figure 5 Representation of NLG vs Visual Synthesis comparison

Figure 5 depicts the relative performance of NLG and Visual Synthesis approach in four assessment measures. The findings indicate that Visual Synthesis scored more in conceptual understanding, creativity and retention, but NLG scored more in narrative coherence, showing the strength of the two complements.

5. CONCLUSION

It had been proven in the study that the use of AI-based digital storytelling was an effective method of redefining the visual narratives in the education sector through the integration of artificial intelligence and creative pedagogy. The results were found to show huge differences between creativity, adaptability, engagement, and retention of learning through the comparative analysis of three AI-based solutions namely: Natural Language Generation (NLG), Generative Visual Synthesis, and Multimodal Emotion and Engagement Analysis. The NLG technique had performed well in narrative coherence, and it is able to give contextually consistent and linguistically rich stories. The method of Visual Synthesis had improved the conceptual knowledge and innovativeness by using realistic and context compatible visualizations. Nevertheless, Emotion and Engagement Analysis approach had proven the most useful with the highest level of learner engagement and understanding and performance improvement of about 28 percent on average compared to others. The results had highlighted adaptive feedback mechanism and emotional intelligence as significant in making digital storytelling an interactive and learner-centered process. AI storytelling is a valid educational innovation because the inclusion of affective computing into storytelling systems had increased motivation, empathy, and long-term retention in the learners.

To expand on the research in the future, it might be considered adding AR/VR technologies, real-time affective sensors and multilingual AI models to design even more immersive and inclusive learning spaces. Also, creating ethical frameworks to provide transparency, data privacy, and bias reduction in AI-driven storytelling systems would further make them responsible and sustainable in the educational field.

CONFLICT OF INTERESTS

None.

ACKNOWLEDGMENTS

None.

REFERENCES

Ginting, D., Woods, R. M., Barella, Y., Limanta, L. S., Madkur, A., and How, H. E. (2024). The Effects of Digital Storytelling on the Retention and Transferability of Student Knowledge. SAGE Open, 14(3). https://doi.org/10.1177/21582440241271267

Gupta, N., et al. (2024). AI-Driven Digital Narratives: Revolutionizing Storytelling in Contemporary English Literature through Virtual Reality. In 2024 Second International Conference on Computational and Characterization Techniques in Engineering and Sciences (IC3TES) (pp. 1–5). IEEE. https://doi.org/10.1109/IC3TES62412.2024.10877467

Hedaoo, M., Jambhulkar, S., Tagade, K., Mate, G., and Durugkar, P. (2025). A Comparative Study of Education Loan Scheme of Punjab National Bank and ICICI Bank, Nagpur 2021–2024. International Journal of Research and Development in Management Review, 14(1), 67–71.

Kumbhar, A. S., Khot, A. S., and Kharade, K. G. (2025). Sentiment Analysis of Amazon Product Reviews. IJEECS, 14(1), 20–25.

Li, Y., et al. (2024). In-Situ Mode: Generative AI-Driven Characters Transforming Art Engagement through Anthropomorphic Narratives. arXiv preprint arXiv:2409.15769.

Naik, I., Naik, D., and Naik, N. (2023). Chat Generative Pre-Trained Transformer (ChatGPT): Comprehending its Operational Structure, AI Techniques, Working, Features and Limitations. In 2023 IEEE International Conference on ICT in Business Industry and Government (ICTBIG) (pp. 1–9). IEEE. https://doi.org/10.1109/ICTBIG59752.2023.10456201

Naik, I., Naik, D., and Naik, N. (2023). Chat Generative Pre-Trained Transformer (ChatGPT): Comprehending its Operational Structure, AI Techniques, Working, Features and Limitations. In 2023 IEEE International Conference on ICT in Business Industry and Government (ICTBIG) (pp. 1–9). IEEE. https://doi.org/10.1109/ICTBIG59752.2023.10456201

Nik, E., Gauci, R., Ross, B., and Tedeschi, J. (2024). Exploring the Potential of Digital Storytelling in a Widening Participation Context. Educational Research, 66(3), 329–346. https://doi.org/10.1080/00131881.2024.2362336

Nik, E., Gauci, R., Ross, B., and Tedeschi, J. (2024). Exploring the Potential of Digital Storytelling in a Widening Participation Context. Educational Research, 66(3), 329–346. https://doi.org/10.1080/00131881.2024.2362336

Preetam, Muppalla, S. C. M., Raj, A., and Chawla, J. (2024). AI Narratives: Bridging Visual Content and Linguistic Expression. In 2024 IEEE International Conference on Smart Power Control and Renewable Energy (ICSPCRE) (pp. 1–6). IEEE. https://doi.org/10.1109/ICSPCRE62303.2024.10675203

Preetam, S. C., Muppalla, A., Raj, A., and Chawla, J. (2024). AI Narratives : Bridging Visual Content and Linguistic Expression. In 2024 IEEE International Conference on Smart Power Control and Renewable Energy (ICSPCRE) (pp. 1–6). IEEE. https://doi.org/10.1109/ICSPCRE62303.2024.10675203

Quecan, L. (2021). Visual Aids Make a Big Impact on ESL Students: A Guidebook for ESL teachers.

Sato, T., Lai, Y., and Burden, T. (2022). The Role of Individual Factors in L2 Vocabulary Learning with Cognitive-Linguistics-Based Static and Dynamic Visual Aids. ReCALL, 34(2), 201–217. https://doi.org/10.1017/S0958344021000288

Yu, W., Peng, J., Li, J., Deng, Y., and Li, H. (2022). Application Research of Image Feature Recognition Algorithm in Visual Image Recognition. In 2022 IEEE Conference on Telecommunications, Optics and Computer Science (TOCS) (pp. 812–816). IEEE. https://doi.org/10.1109/TOCS56154.2022.10016132

Yu, W., Peng, J., Li, J., Deng, Y., and Li, H. (2022). Application Research of Image Feature Recognition Algorithm in Visual Image Recognition. In 2022 IEEE Conference on Telecommunications, Optics and Computer Science (TOCS) (pp. 812–816). IEEE. https://doi.org/10.1109/TOCS56154.2022.10016132

Zhang, Y., et al. (2025). AI-Driven Optimization Algorithm for Cultural Dissemination Pathways and its Educational Applications. In Proceedings of the 2025 International Conference on Artificial Intelligence and Educational Systems. Association for Computing Machinery. https://doi.org/10.1145/3744367.3744429