

The Power of Visual Storytelling: A Deep Learning Framework for Educational Influence

K. Dinesh Kumar^{*1}, Sarot Srang², Dona Valy³

Submitted: 27/12/2023 Revised: 03/02/2024 Accepted: 11/02/2024

Abstract: This research focuses on enhancing students' memory retention and visualization skills in education through visual storytelling, a powerful pedagogical approach known for improving learning outcomes across various domains. Our visual storytelling method involves generating images from text using Generative Adversarial Networks (GANs) along with the Graph Embedding technique, contributing to comprehension, memory, sequencing, and engagement. A primary survey revealed that a significant portion of teachers (70.3%) in India and Cambodia incorporate visual storytelling in half of their total class time. However, the process is perceived as time-consuming, prompting educators' expectations for a unified model capable of generating complete reading comprehension along with images. In our post-experiment test, involving 8th-grade students, the study compared the Experimental Group, experiencing visual storytelling, with the Control Group using traditional methods. Assessments through Fill-in-the-Blanks, Multiple Choice Questions (MCQs), and story continuation demonstrated the Experimental Group's improved recall, application, and critical thinking skills. Remarkably, 92% of the Experimental Group actively engaged in continuing the story, signifying a substantial increase compared to the Control Group. Post-research surveys show a strong willingness of educators (96.3%) to integrate visual storytelling into classrooms, with 63% expressing a desire to incorporate our research findings into teaching methods, particularly in reading comprehension with image sequences. These findings underscore the positive impact of visual storytelling, advocating for its widespread adoption and further exploration in educational practices.

Keywords: GANs, Nodes, Relation, Graph Embedding, Visual Storytelling, Blooms Taxonomy

1. Introduction

From ancient times to the present day, storytelling has served as one of the most significant educational tools, enabling the sharing of knowledge from one generation to the next. Many of us cherish the tales and stories shared by our grandparents during holidays. Even now, these stories are like beautiful pictures woven into our memories of growing up. The power of these stories is remarkable. As we listened, whether they were factual or fictional, they could transport us into different situations, igniting our imaginations and leading us on adventures beyond the boundaries of our everyday lives. Times change, but the memories of these stories help us imagine how life was in their youth, the feelings they experienced, the adventures they had, and their expectations. The ability of storytelling to make concepts tangible and visually appealing can revolutionize education, allowing students to grasp and internalize complex ideas with greater depth and clarity.

Storytelling proves to be a powerful pedagogical approach that can improve learning outcomes across various domains, including general, scientific, and technical education [1]. Children enjoy reading, writing, listening to, and watching stories. Through these stories, they gain a deeper understanding of society and life in general. Both storytelling and story reading have been found to be successful educational strategies that lead to significant improvements in language acquisition for young children, enhance oral language skills, develop reading comprehension, explain scientific concepts, facilitate effective communication, improve cross-cultural understanding, and promote moral and social development [1]. Furthermore, storytelling has undergone a remarkable evolution, adapting to modern times and embracing diverse forms and applications. The rise of digital technologies has revolutionized the art of storytelling, introducing innovative collaborative approaches. Fuelled by technological advancements, digital storytelling methods have witnessed widespread adoption due to their simplicity, interactivity, and affordability. Digital storytelling finds application across all education levels, from early childhood to higher education. Many educators are incorporating digital technology into their storytelling practices, and an increasing number of schools are adopting digital storytelling techniques that blend the art of storytelling with various digital multimedia elements, including images, audio, and video.

¹ *Mechatronics and Information Technology, Institute of Technology of Cambodia, Phnom Penh, Cambodia*
ORCID ID: 0009-0001-5586-7723

² *Mechatronics and Information Technology, Institute of Technology of Cambodia, Phnom Penh, Cambodia.*
ORCID ID: 0000-0001-6917-6352

³ *Department of Information and Communication Engineering, Institute of Technology of Cambodia, Phnom Penh, Cambodia*
ORCID ID: 0009-0000-6181-5551

* Corresponding Author Email: vkjdinesh@gmail.com
vkjdinesh@gmail.com, srangsarot@itc.edu.kh, dona@itc.edu.kh

* Corresponding Author Email: author@email.com

The integration of technology into classrooms by educational practitioners offers children enhanced learning possibilities. As a result, children are entitled to equal opportunities to participate in and derive benefits from these educational advancements. Digital storytelling, in addition, plays a crucial role in nurturing students' creativity, enabling them to address important issues in inventive ways. Serving as an effective pedagogical tool, digital storytelling enhances learner motivation and contributes to a conducive learning environment. It facilitates collaborative story construction through reflection and interpersonal communication. Students can utilize multimedia software tools and employ various technology skills to create digital stories focused on specific educational issues [2]. Various technological advancements and Edutech software play an important role in facilitating digitized education within the education sector. This includes storytelling through subscription models tailored for business perspectives. However, many of these tools are not readily accessible to a wide audience. When we think of storytelling, the common association involves generating text from images, a prevalent practice in education. In primary education, textbooks often feature vibrant pictures with attractive colors to elucidate concepts and capture students' attention. However, as education progresses beyond the primary level, there are fewer visual aids in the textbook. Our approach aims to bridge this gap.

The main objective of this study is to improve students' memory retention and visualization skills by creating engaging images from textual content. This methodology enhances visualization skills by encouraging the articulation of ideas, emotions, and inquiries. The act of reading stories helps children associate words with meaning, motivating them to enhance their reading and writing abilities. Visual storytelling plays a crucial role in promoting comprehension, memory retention, sequencing, and overall engagement. Moreover, it nurtures a child's imagination and creativity, broadening their perspective [3]. We, as humans, are strongly influenced by visuals, and information presented visually is more effectively processed and retained [4]. Visuals possess the ability to simplify complex concepts, making it easy for users with different abilities to quickly understand the intended message. The goal of visual storytelling, in general, is to establish an emotional connection with the audience by utilizing imagery that resonates with them [4]. Our working memory comprises both verbal and visuospatial components, and their effective integration plays a pivotal role in enhancing overall understanding. To put it simply, the synergy between images and text can significantly improve comprehension. Without this collaborative approach, there is a risk of overloading our working memory, potentially leading to a decline in performance. The incorporation of images not only serves to augment

learning but also facilitates mental visualization, contributing to enriched reading comprehension.

The human brain possesses the remarkable ability to comprehend entire images in as little as 13 milliseconds. This rapid processing speed allows us to focus on various elements, with our eyes shifting focus three times per second. Recent research indicates that approximately 65% of individuals exhibit improved learning outcomes when presented with images, surpassing the effectiveness of text and speech alone [5]. Images possess the unique capability to convey complex ideas swiftly and evoke powerful emotions that extensive explanations may struggle to match. Remarkably, about 75% of the information processed by our brains is visual [6]. This underscores the importance of incorporating visual elements for better understanding and retention.

In the realm of education, these insights hold significant implications. Students stand to benefit from a variety of visual aids, including text-to-images, videos, graphs, cartoons, games, and flashcards. Both educators and students can leverage these visuals to enhance comprehension, establish connections between different textual elements, and create a more engaging learning experience.

2. Methodology

This research aims to improve students' memory retention and visualization skills in education through the implementation of visual storytelling. We generate a story from text prompts, allowing users to click on each sentence to generate corresponding images. Additionally, we create Multiple Choice Questions (MCQs) and fill-in-the-blanks derived from the visual storytelling process to assess students' memory retention and visualization skills. The entire process is facilitated through a dedicated web application developed for this purpose.

The research comprises several phases. In the initial stage Preliminary Data Collection Survey: Gathering Insights for Research, 2nd Stage : Structuring Visual Storytelling: Emphasizing Reading Comprehension, 3rd Stage :Generating Short Stories from Prompts, 4th Stage: Sentence Segmentation, 5th Stage: Extracting Noun Phrases and Relations, 6th Stage: Constructing Directed Graph and Graph Embedding, 7th Stage: Generating Images Using Generative Adversarial Networks (GANs), 8th Stage: Generating Multiple-Choice Questions (MCQs), 9th Stage: Creating Fill in the Blanks, 10th Stage: Web Application with Customized GenAI for Visual Storytelling, 11th Stage: Experiment Conducted with Group Students to Assess Effectiveness, Feedback from Students and Teachers.

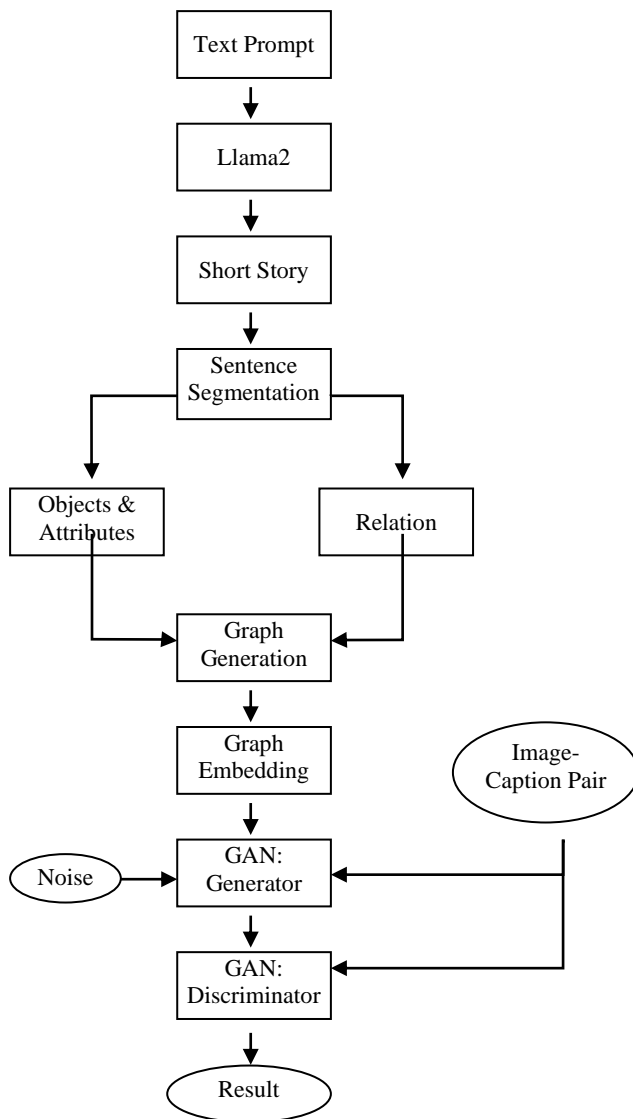


Fig 1. Model Architecture

Following the implementation of visual storytelling, we introduced it to students and teachers in a middle school setting, specifically targeting students aged 11 to 13 years old. We organized two study groups among the students. The control group engaged in traditional reading activities, answering Multiple Choice Questions (MCQs), Fill in the Blanks, and continuing the story. In contrast, the experimental group utilized the complete model, integrating images into the learning process. We assessed their performance based on Bloom's taxonomy standards of remembering, understanding, and applying. Consequently, the experimental group outperformed the control group, with 92% of students actively contributing to the continuation of the story. This indicates a substantial 52% improvement compared to the performance of the control group.

2.1. Preliminary Data Collection Survey: Gathering Insights for Research

We initiated the preliminary phase of data collection by engaging with primary and middle school teachers in India

and Cambodia. The primary objective was to evaluate the potential impact of incorporating visual storytelling techniques into educational practices. Our focus centered on identifying practical and effective methods for seamlessly introducing visual storytelling within the classroom environment, particularly in the transition from textual content to engaging visual representations. The study also aimed to delineate specific activities associated with this integration to provide valuable insights into the feasibility and efficacy of visual storytelling as an educational tool.

In the survey, 77.8% of respondents identified as female, while 22.3% identified as male. Most responses (74.1%) originated from India, with 25.9% coming from Cambodia. Regarding language proficiency, 96.3% of participants believed that primary school students could comprehend English well enough to benefit from visual storytelling. In terms of its impact, 51.8% asserted that visual storytelling enhances students' visualization skills, while only 7.4% disagreed.

A significant portion of teachers (70.3%) reported incorporating visual storytelling in half of their total class time. Of these, 67% found visual storytelling to be effective in language learning, and nearly all observed improvements in students' memory retention. Considering the approach, 52% of educators were familiar with visual storytelling from text to image. A noteworthy 63% considered the combination of text and image as an effective learning method.

Encouragingly, 96.3% expressed willingness to integrate visual storytelling into their classes, with 63% strongly desiring to incorporate the research findings into their teaching methods such as reading comprehension along with image sequences. Additionally, 11.1% expressed a moderate willingness to do so, highlighting overall positive mindset toward the proposed visual storytelling educational approach.

2.2. Structuring Visual Storytelling: Emphasizing Reading Comprehension

As a reflection of the preliminary data collection to improve students' memory retention and visualization, reading comprehension is crucial for the ability to read, process, and comprehend the meaning of text/story. This involves developing two key skills: word reading and language comprehension. Our research goes beyond mere memorization of words and phrases; instead, we aim to cultivate a mental model of the story by integrating the meaning of words and sentences, reinforced with relevant images for each sentence. Our research aims to cultivate strong reading comprehension skills, ensuring that readers not only understand but also enjoy what they are reading

[7]. This involves incorporating images as part of visual storytelling.

We chose to apply our research to contribute to the education sector. Our approach involves taking text prompts from users to generate short stories. We extract sentences one by one to produce relevant images that aid in conveying the concepts to readers. To assess students' understanding, we created Multiple Choice Questions (MCQs) and fill-in-the-blanks derived from the generated story, enabling the evaluation of students' memory retention skills. Additionally, students were prompted to continue the story, providing insights into their visualization capabilities.

2.3. Generating Short Stories from Prompts

Today's educators face various challenges, especially in the post-pandemic education landscape. Their goal is to help students understand the complexities of today's world and prepare them for the uncertainties of tomorrow [8]. Consequently, teachers must be attuned to effective teaching methods for this era, understand students' needs, and continually seek improvement. Addressing these challenges, digital technologies play a crucial role in significantly enhancing teaching and learning in primary and secondary schools. Bringing the latest technology into the classroom will consistently fulfil the goal of instigating change.

Creating visual stories enhances students' memory retention and visualization skills. The use of advanced language models, like Llama2, facilitates the generation of short stories based on given text prompts. This process improves students' comprehension, enabling them to recognize the connection between the generated story and the prompt. It also enhances their ability to predict and make decisions about the story's progression, contributing to improved memory retention and visualization skills.

The Large Language Model Llama 2, which was recently introduced by Meta and is open source, was used in our research. Llama 2 pretrained models were trained on 2 trillion tokens, providing double the context length compared to Llama 1[9]. Its fine-tuned models were trained with over 1 million human annotations, enabling us to generate short stories based on prompts. The number of sentences in the generated story was identified, and segmentation was performed based on the periods between sentences. The index was generated for the sentence in the story, creating links to generate images by clicking on each sentence.

2.4. Extracting Noun Phrases and Relations

Our research aims to create images based on textual descriptions while ensuring semantic coherence among the objects, their attributes, and the relationships mentioned in

the text. To achieve this, we incorporate natural language processing (NLP) techniques, specifically integrating part-of-speech tags, building upon the foundation of our base paper [10]. Objects and their attributes are identified as nodes using the following grammar:

Nodes: {<DT|JJ|JR|JJS|RB|RBR|RBS|CD>*<NN. *>+}

This grammar includes determiners, adjectives, adverbs, and nouns, allowing us to extract meaningful nodes for maintaining semantic consistency.

Relation: {<IN|VBG|VBN|VBP|VBZ>}

In this process, we focus on capturing prepositions and verbs that appear between the nodes in the sentence, serving as edges or relations. Through this approach, we extend the objectives outlined in our base paper [10].

2.5. Constructing Directed Graph and Graph Embedding

While constructing the graph for the text input, we utilize sentence segmentation to extract nodes and relations from sentences, forming the basis of the graph. For each sentence, a sequence graph is generated by establishing connections between various nodes based on their relationships. Considering that a sentence may encompass multiple nodes, the graph structure adheres to the "Source Node: Relation: Destination Node" format, preserving the sequential order of the sentence [8]. Once the graph for a sentence is established, the subsequent step involves graph embedding. From the array of graph embedding techniques available, we opt for the Node2Vec approach, known for providing graph.

2.6. Generating Images Using Generative Adversarial Networks (GANs)

In the GAN framework, training involves both the Generator and Discriminator utilizing image-level embeddings from our dataset. This entails using node embeddings extracted from provided text descriptions along with random noise as inputs to the GAN Generator [10]. The outcome is the creation of synthetic images that are then evaluated and categorized by the Discriminator.

After the images have been generated for the text description, as our goal is to enhance the memory retention and visualization of the students, we assess their performance by conducting tests such as Multiple-Choice Questions and Fill-in-the-blanks which are generated for the story. Assessment of students' memory retention is categorized according to Bloom's Standard in the following ways: Test 1: Fill in the blanks - Assessing Remembering Skills, Test 2: Multiple-Choice Questions (MCQs) - Evaluating Understanding Skills.

Table 1. Bloom’s Taxonomy based experiments/Test

Experiments/Test	Skills	Bloom's Taxonomy Levels
Fill in the blanks		Remember (L1)
Multiple Choice Questions	Memory Retention	Understand (L2)
Continue the story	Visualization	Apply (L3)

The objective of our research is to create comprehensive reading comprehension materials, complete with relevant images for each sentence to improve the student's memory retention and visualization skills. Additionally, we aim to provide assessments that will assist teachers in preparing lessons for their classes, enabling students to engage in effective practice. The MCQs will be generated using the open-source LLM, such as Llama 2. We created a "Fill in the Blanks" activity using Part of Speech Tags for the generated story and identified the prepositions in each sentence, left them empty, and prompted students to fill in the blanks.

3. Result and Discussion

To assess the effectiveness of our research, we conducted tests with 8th-grade (Age Range between 11 to 13) students. The study group was divided into two categories: the Control Group and the Experimental Group. In the Control Group, students were given the generated short story and instructed to read independently for 15 minutes, followed by a 30-minute examination. On the other hand, students in the Experimental Group were provided with the generated short story accompanied by images for each sentence. They, too, were instructed to read independently for 15 minutes, followed by a 30-minute examination. The 30-minute exams include a combination of Multiple-Choice Questions and Fill in the Blanks. These elements are designed to evaluate students' memory retention. Additionally, students are prompted to continue the story, providing an assessment of their visualization skills.

We gathered a total of 424 samples, with 212 samples in each group. As mentioned earlier, the control group adhered to traditional reading comprehension methods, including assessments. In contrast, the experimental group followed the same procedures as the control group, with the additional enhancement of having generated images accompanying each sentence in the story. The research hypothesis suggests that the performance of students in the experimental group should surpass that of the control group. The result of the experiment gave us valuable insights about introducing visual storytelling in educational practice.

3.1. Memory Retention Task 1: Fill in the Blanks

Table 2. Result of Fill in the Blanks exercise

Control Group		Experimental Group	
Score Range	# Students	Max Range	# Students
0	8	0	3
1 - 5	64	1 - 5	53
6 - 10	57	6 - 10	48
11 - 15	27	11 - 15	28
16 - 20	14	16 - 20	23
21 - 25	42	21 - 25	57
More	0	More	0

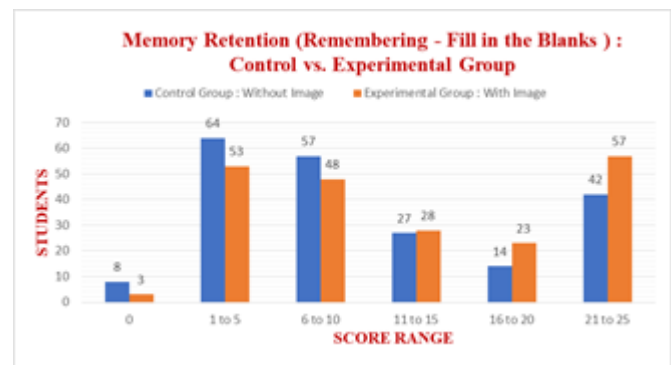


Fig 2. Graphical representation of outcome of the fill-in-the-blanks exercise

In the zero scores (0), both groups exhibit a decrease in the count of participants scoring zero in the experimental group, indicating an improvement in basic understanding at the memory recall level. Moving to low scores (5 and 10), the experimental group demonstrates a reduction in the number of participants scoring 5 and 10 compared to the control group, suggesting that the inclusion of relevant images for each sentence contributes to better memory retention at lower difficulty levels within Bloom's Level 1[11]. At moderate scores (15), the count of participants scoring 15 remains similar in both groups, implying that the introduction of visual elements may have a more substantial impact on scores at lower difficulty levels, aligning with the memory retention objectives of Bloom's Level 1. Transitioning to higher scores (20 and 25), the experimental group exhibits an increase in participants scoring 20 and 25, indicating enhanced memory retention and comprehension skills at higher difficulty levels within Bloom's Level 1 when images accompany the text. Regarding outliers (More), both groups show no participants in the "More" category, suggesting a consistent upper limit in scores within the studied sample. Overall, the experimental group's performance demonstrates an encouraging trend, aligning with the objectives of Bloom's Level 1 activities. Fewer participants scored in the lower ranges, indicating that including relevant images in visual storytelling positively influences students' memory

retention and comprehension skills at the foundational level of Bloom's Taxonomy.

3.2. Memory Retention Task 2: MCQs

Table 3. Result of Fill in the Blanks exercise

Control Group		Experimental Group	
Score	# Students	Score	# Students
0	8	0	2
1	20	1	20
2	28	2	30
3	46	3	47
4	60	4	58
5	50	5	55

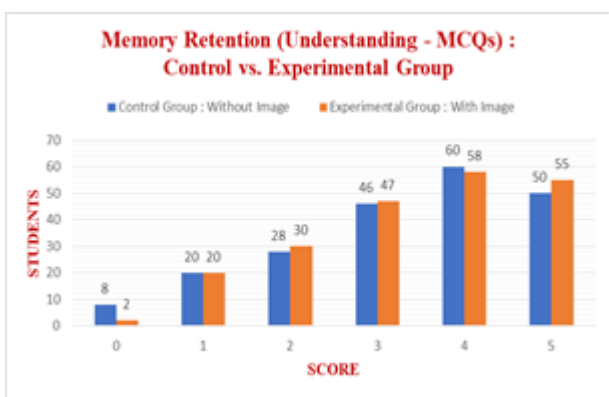


Fig 3. Graphical representation of outcome of the Multiple Choice Questions exercise

In evaluating students' understanding at Bloom's Level 2[11], focusing on comprehension and application skills, a test comprising Multiple-Choice Questions (MCQs) was administered to both control and experimental groups. The collected data, consisting of 212 samples in each group, reveals notable insights. The control group exhibited varied scores with the majority falling within the moderate range, while the experimental group showcased a discernible shift. Notably, the experimental group demonstrated a decrease in zero scores, indicating enhanced recall. Moreover, participants in the experimental group showed improved performance in the higher score range (5), suggesting a positive correlation between the inclusion of relevant images in visual storytelling and students' application and critical thinking skills. The slight reduction in lower scores in the experimental group implies an encouraging impact on basic comprehension at Bloom's Level 2. The absence of outliers in both groups suggests a consistent upper limit in the scores within the studied sample. This analysis underscores the potential benefits of integrating

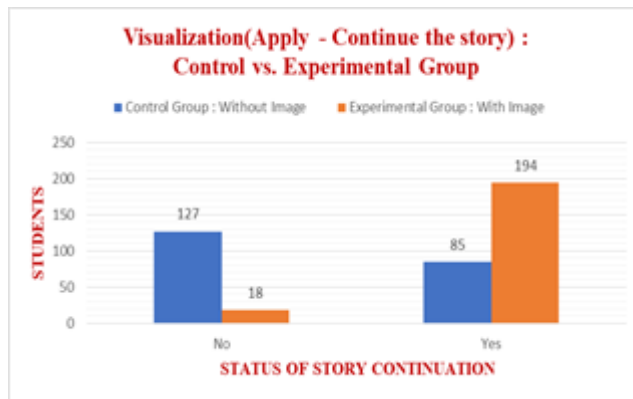


Fig 4. Graphical representation of outcome of the story continuation exercise

visual elements into educational practices, fostering improved comprehension and application of knowledge among students.

3.3. Visualization Skill Task 3: Story Continuation

To assess students' visualization skills aligned with Bloom's Level 3, an assignment was provided, tasking them with continuing the story they had read. Performance evaluation was based on the enthusiasm demonstrated in their story continuation. Visualization assessment was categorized into yes/no responses. In the control group, 40% of the students attempted to continue the story. In contrast, the experimental group exhibited a noteworthy difference, with an impressive 92% of the students actively engaging in the task of continuing the story. This substantial increase in participation within the experimental group suggests a positive impact on students' visualization skills when relevant images accompany the text, supporting the objectives of Bloom's Level 3.

In summary, the findings indicate that incorporating visual elements into educational practices, specifically relevant images in visual storytelling, holds promise for improving students' memory retention, comprehension, and visualization skills across different cognitive levels. This research contributes valuable insights for educators seeking effective strategies to enhance students' learning experiences.

4. Conclusion

This research highlights the transformative impact of visual storytelling on students' memory retention, comprehension, and visualization skills. By employing deep learning approaches and a comprehensive research methodology, the study reveals a positive trend in the experimental group's performance, showcasing improved memory recall, comprehension, and visualization skills. The integration of visual storytelling not only led to decreased lower scores but also resulted in increased higher scores, aligning with objectives at Bloom's Level 1 and Level 2. Furthermore, it had a positive impact on

students' visualization skills, as demonstrated by their story continuation performance under Bloom's Level 3. Specifically, 92% of the students actively engaged in the task of continuing the story. The findings, supported by educator insights and preliminary surveys, strongly advocate for the widespread adoption of visual storytelling in education. This approach offers a dynamic and engaging method to enhance learning outcomes. Overall, this research provides valuable insights, opening avenues for further exploration and emphasizing the pivotal role of technology in creating enriched and interactive learning environments.

References

- [1] Maila D. H. Rahiem “Storytelling in early childhood education: Time to go digital”, *International Journal of Child Care and Education Policy, ICEP* 15, 4 (2021). <https://doi.org/10.1186/s40723-021-00081-x>.
- [2] Smeda, N., Dakich, E. & Sharda, N. “The effectiveness of digital storytelling in the classrooms: a comprehensive study”. *Smart Learn. Environ.* 1, 6 (2014). <https://doi.org/10.1186/s40561-014-0006-3>
- [3] Sinead McGee,” The role of storytelling in early literacy development at Northbridge”, *The role of storytelling in early literacy development at Northbridge (nordangliaeducation.com)*, 26 April, 2021.
- [4] The power of visual storytelling in content design (uxdesigninstitute.com)
- [5] Vinnie Wong, “What is visual Storytelling? How to engage and Inspire Audience”, Aug 31, 2023, *What Is Visual Storytelling? How to Engage and Inspire Audiences (piktochart.com)*.
- [6] Jamal Raiyn, “The Role of Visual Learning in Improving Students’ High-Order Thinking Skills”, *Journal of Education and Practice*, www.iiste.org ISSN 2222-1735 (Paper) ISSN 2222-288X , (Online) Vol.7, No.24, 2016.
- [7] What is reading comprehension and why is it important? (oxfordowl.co.uk)
- [8] Halverson, R., & Shapiro, R. B. (2012). *Technologies for education and technologies for learners: How information technologies are (and should be) changing schools.* (WCER Working Paper no. 2012-6).
- [9] Llama 2 - Meta AI
- [10] K. Dinesh Kumar, Dr. Sarot Srang, Dr. Dona Vally, “Visual Storytelling: A Generative Adversarial Networks (GANs) and Graph Embedding Framework”, *International Journal on Recent and Innovation Trends in Computing and Communication*, 30-September-2023.
- [11] Christine Persaud, “Bloom’s Taxonomy: The Ultimate Guide”, *HIGHER ED TRENDS*, *How to Implement Bloom's Taxonomy in Your Course | Top Hat*, February 25, 2023.