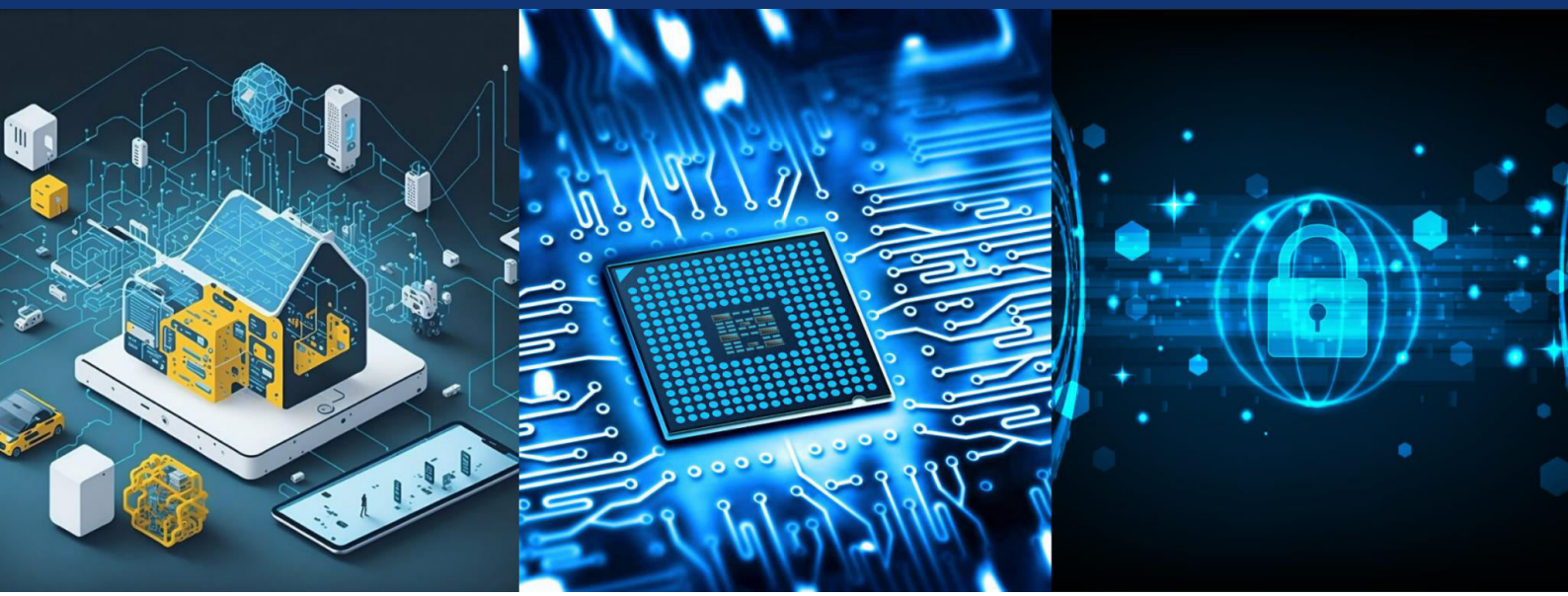
IJIRCCE

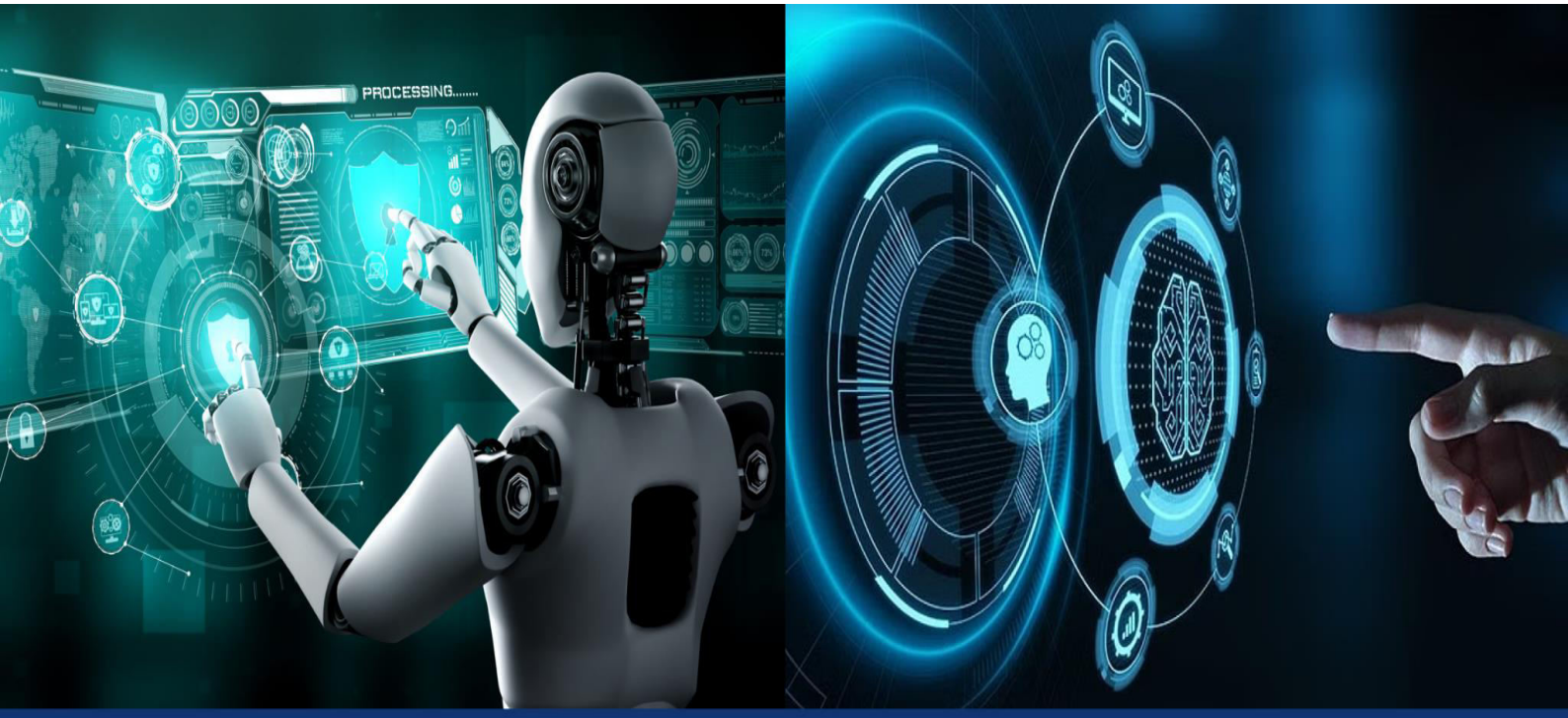
ISSN(O): 2320-9801

ISSN(P): 2320-9798



International Journal of Innovative Research in Computer and Communication Engineering

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)



Impact Factor: 8.771

Volume 13, Issue 5, May 2025

© 2025 IJIRCCE | Volume 13, Issue 5, May 2025| DOI:10.15680/IJIRCCE.2025.1305072

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.771| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Pixel Puranas: History Revived using GenAI

Dr.Gangadharappa S [1], Chirantha K A [2], Hiranmayi K M [2], Jagruthi B J [2], Vijetha N[2]

Associate Professor, Department of CSE, Bapuji Institute of Engineering and Technology, Davangere, Karnataka, India[1]

UG Student, Department of Computer Science and Engineering, Bapuji Institute of Engineering and Technology, Davangere, Karnataka, India[2]

ABSTRACT: The preservation and presentation of historical and cultural heritage are often constrained by the availability of physical artifacts and the limitations of traditional archival methods. This project, Pixel Puranas: History Revived Using GenAI, leverages generative AI to recreate and reimagine historical narratives, artwork, and architecture through digital media. By integrating machine learning models, particularly image generation and natural language processing, the project reconstructs lost or damaged cultural elements and presents them in an interactive and accessible format. The aim is to bridge the gap between modern technology and ancient heritage, providing immersive educational experiences and promoting cultural awareness. This paper outlines the design, implementation, and societal impact of the system, showcasing how AI can be a powerful tool in historical storytelling and digital preservation.

KEYWORDS*:* Generative AI, Cultural Heritage, Digital Preservation, Historical Reconstruction, Pixel Puranas, AI in Art, Deep Learning, Interactive History, Image Generation, NLP in History.

I. INTRODUCTION

Understanding history is a cornerstone of cultural education, yet students often struggle with the abstract and complex nature of historical narratives, events, and their significance.

Traditional teaching methods, which primarily rely on textual explanations and static images, fail to engage students in an immersive and interactive manner, limiting their ability to truly connect with the content. The increasing demand for more engaging and dynamic educational tools highlights the need for innovative solutions that bridge the gap between theoretical knowledge and practical, visual understanding.

Generative Artificial Intelligence (AI) has emerged as a transformative tool in education, offering new ways to present complex topics through dynamic and interactive visualization. Recent advancements in AI, particularly in text-to-image generation and multimedia synthesis, emphasize the potential of AI in creating engaging and personalized learning experiences.

To address these challenges, we propose Pixel Puranas: History Revived, an AI-powered platform that uses Generative AI to enhance the teaching of Indian history. This project, developed using GPT-Neo for text generation and Stable Diffusion for image creation, enables students to visualize historical narratives through AI- generated images and videos. The system not only generates detailed historical stories but also combines these images and stories into a vido format with audio narration, making history more engaging and accessible. This paper presents the design and implementation of Pixel Puranas, demonstrating how AI and multimedia can transform the way history is taught and experienced, offering a dynamic and interactive learning process. By combining AI-generated content with multimedia storytelling, the proposed system aims to create a deeper connection with historical knowledge, fostering better understanding and retention for modern learners.

II. DATASET

For this research, to ensure a comprehensive and representative training and testing environment, we utilized a diverse collection of publicly available and curated datasets. These included the WikiArt dataset for stylistic references, the

Indian Heritage Dataset for culturally significant visuals, and classical Indian literary sources such as the Vedas, Puranas, Mahabharata, and Ramayana. This rich combination of visual and textual data served as a foundational base for training our generative AI models, enabling accurate and immersive historical reconstructions in both visual storytelling and narrative synthesis.The Indian Heritage Dataset contains a wide range of culturally rich and historically significant images, capturing real-world representations of monuments, artifacts, and traditional art forms. It offers authentic visual examplesessential for historical accuracy. Additionally, the WikiArt dataset contributes diverse artistic styles that enhance the visual quality and depth of generative outputs. We further incorporated classical Indian textual sources such as the Vedas, Puranas, and epics like the Ramayana and Mahabharata. By integrating these visual and narrative resources, we constructed a comprehensive dataset comprising both historical context and artistic interpretation, allowing our generative AI models to learn and generate coherent, culturally grounded visual-narrative content.

III. RELATED WORK

Many research papers have contributed to the development of multimodal generative systems combining natural language processing and deep generative models. Each of these works offers insights into solving challenges such as visual consistency, historical accuracy, narrative coherence, and user interaction. Below are key contributions in this field that have influenced the direction of our project.

(Vivian Liu, 2023)*:* This study explores multimodal prompts in generative AI systems, emphasizing their importance in enhancing user creativity and control. The research introduces structured workflows that help users translate abstract goals into visually and narratively coherent outputs. The findings align with our use of prompt engineering in Pixel Puranas to convert historical text into image sequences that follow logical, artistic, and cultural structure.

(Chenshuang Zhang et al., 2023)*:* In their survey on text-to-image diffusion models the authors provide an in-depth analysis of various generative techniques that transform language into visual content. Their work highlights the capabilities and limitations of current diffusion-based models, particularly in rendering historical or culturally nuanced content. Our system builds on these principles by tailoring image generation to Indian heritage themes and textual epics.

(Jay Zhangjie Wu et al., 2023 – Tune-A-Video)*:* This paper introduced an approach to convert a single image into a coherent video sequence using text guidance, enabling a step-by-step evolution of scenes. This directly inspired our module that animates generated historical images into short clips to depict events dynamically, enhancing narrative immersion in Pixel Puranas.

*(*Robin Rombach et al., 2022)*:* Their work on Retrieval-Augmented Diffusion Models presents a way to prompt generative models using visual references, not just text. This was particularly influential in our incorporation of the Indian Heritage Dataset and artistic styles from WikiArt to maintain visual consistency and contextual authenticity in image generation.

(Li Hu et al., 2024 – Animate Anyone)*:* This study introduced a character animation pipeline that ensures temporal consistency and visual fidelity when converting still images into animated sequences. We adapted similar strategies to maintain character integrity and movement smoothness during the animation of mythological figures in our platform.

(Jiahong Su and Weipeng Yang, 2024 – IDEE Framework)*:* Their work proposed a theoretical model for incorporating AI tools into educational settings. This framework provided the backbone for the interactive layer of Pixel Puranas, including quiz-based learning, voice narration, and historical Q&A functionality designed for student engagement.

(Doa Kim et al., 2024)*:* In exploring intent-reflective image generation, this research showed the need for aligning model outputs with user expectations throughout different creative stages. The dynamic prompting and user-controlled storytelling in our system were informed by this approach, allowing customization of visuals and narratives based on user preferences.

(Mouna Rabhi et al., 2024) [11]: This paper examines the Weakness of AI-based audio authentication systems toadversarial attacks using deepfake audio. We demonstrate that state-of-theart audio deepfake classifiers, such as the

© 2025 IJIRCCE | Volume 13, Issue 5, May 2025| DOI:10.15680/IJIRCCE.2025.1305072

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.771| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Deep4SNet model with 98.5% accuracy, can be severely compromised. Our two adversarial attacks, exploiting generative adversarial networks, reduce the accuracy to nearly 0%, especially to 0.08% in a gray-box scenario.
This project builds on existing research by proposing a novel hybrid CNN-Bi-LSTM architecture that focuses on performance and attack for real-world noise optimization. Through the combination of CNN and Bi-LSTM techniques, the method attempts to detect difference and proximity, improve detection value in the noise environment, and prevent adversarial interference.

IV. METHODOLOGY

The proposed system integrates **Generative Artificial Intelligence (AI)**, **Multimedia Technologies**, and **Video Production** to create an immersive and interactive learning platform for Indian history. By leveraging AI models like **GPT-Neo** for text generation and **Stable Diffusion** for image creation, the system aims to offer a dynamic and engaging experience for students, transforming traditional history education into an interactive learning journey.

A. System Overview

The architecture of the proposed system consists of two main modules:

- **Text and Image Generation Module**: This module utilizes **GPT-Neo** for generating historical narratives based on user prompts and **Stable Diffusion** for creating corresponding images. These generated elements are then used to construct a narrative-driven visual experience.
- **Multimedia Presentation Module**: This module processes the generated text and images into a cohesive video format, accompanied by audio derived from the historical text. The video is designed to enhance the understanding of historical events through dynamic, visual storytelling

B. Text and Image Generation Module

The text and image generation module employs **GPT-Neo** and **Stable Diffusion**.

- **Text Generation with GPT-Neo**: The system takes a user-inputted prompt, such as a specific historical event or figure, and generates a detailed text-based narrative describing the event or figure. This allows users to explore historical content interactively.
- **Image Generation with Stable Diffusion**: Based on the generated text, **Stable Diffusion** is used to create AI-generated images that represent key moments, figures, or scenes from history. These images offer a visual representation that helps students visualize the historical context.

C. Multimedia Presentation Module

Once the text and images are generated, the system combines them to create a video format.

- **Video Creation**: The generated images are sequentially placed within the video timeline to illustrate the historical story. The text is then converted to audio, which narrates the historical events, offering an immersive experience. The video synchronizes the visual and auditory components to help reinforce the concepts.
- **Audio Narration**: The audio is created by converting the generated text into voice narration, ensuring the historical content is both informative and engaging.

D. Frontend Integration

To make the system interactive, the generated content is linked to a front-end interface developed with **HTML**and **CSS**. The frontend allows users to input specific prompts to generate historical content dynamically.

- **User Input Interface**: Users enter their historical queries or topics, and the system generates corresponding content, including text, images, and video. This interaction allows for exploration and engagement with different aspects of Indian history.
- **Dynamic Content Presentation**: The frontend presents the generated content in a user-friendly interface, allowing users to explore different historical narratives through video and audio.

E. Workflow of the System

The system follows a step-by-step workflow to ensure a seamless interaction and learning:
User Input: The user provides a prompt related to Indian history (e.g., a historical event or figure).
Text Generation: **GPT-Neo** generates a detailed historical narrative based on the prompt.
Image Generation: **Stable Diffusion** creates corresponding images to visually represent the narrative.
Video and Audio Creation: The generated text is converted into audio, and the images are placed within a video

© 2025 IJIRCCE | Volume 13, Issue 5, May 2025| DOI:10.15680/IJIRCCE.2025.1305072

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.771| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

format to create a dynamic historical story.

Frontend Interaction: The generated video, images, and text are presented to the user through an interactive frontend.

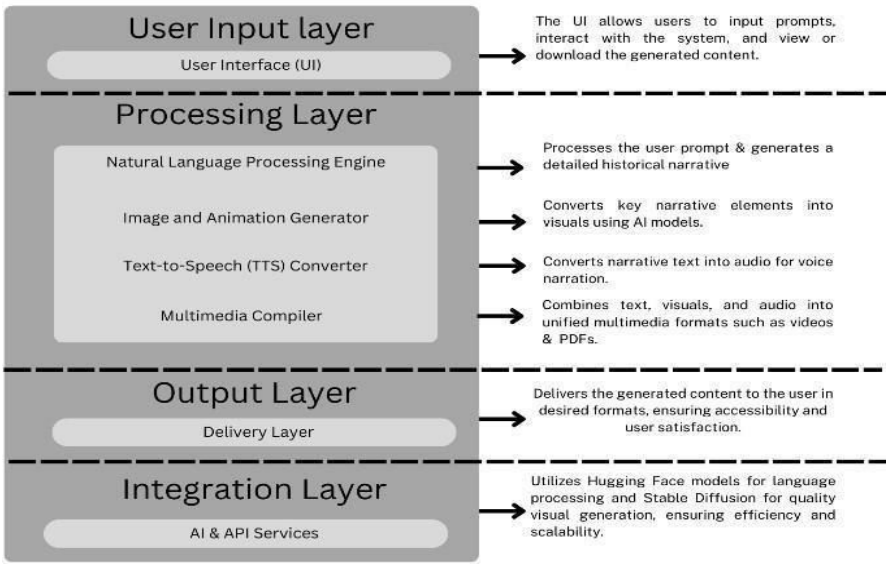


Figure 1: System Architecture

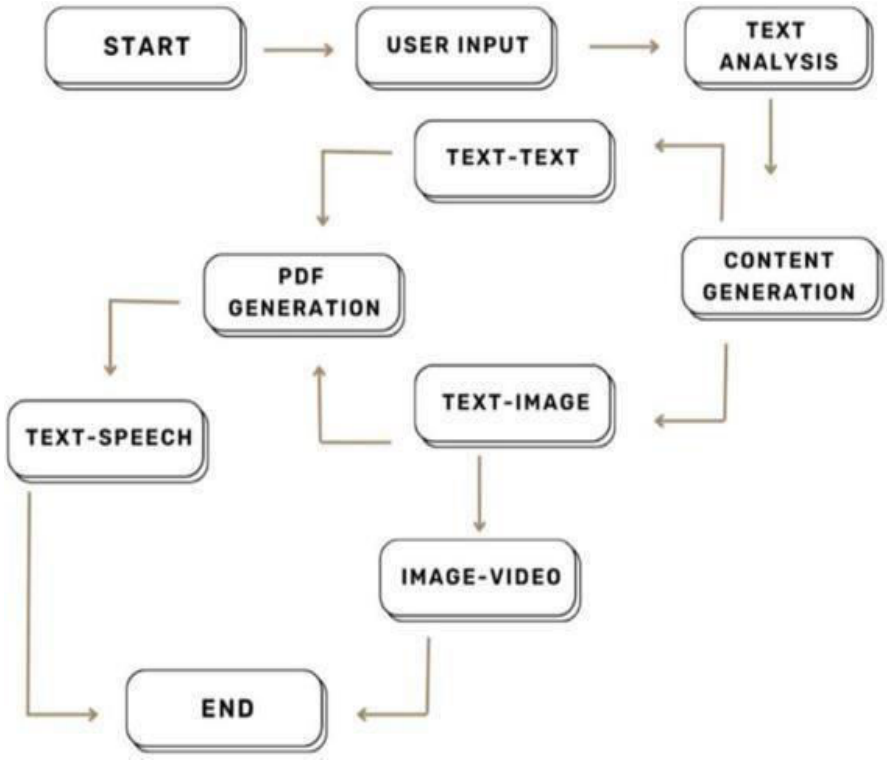


Figure 2: Data Flow Diagram

© 2025 IJIRCCE | Volume 13, Issue 5, May 2025| DOI:10.15680/IJIRCCE.2025.1305072

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.771| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

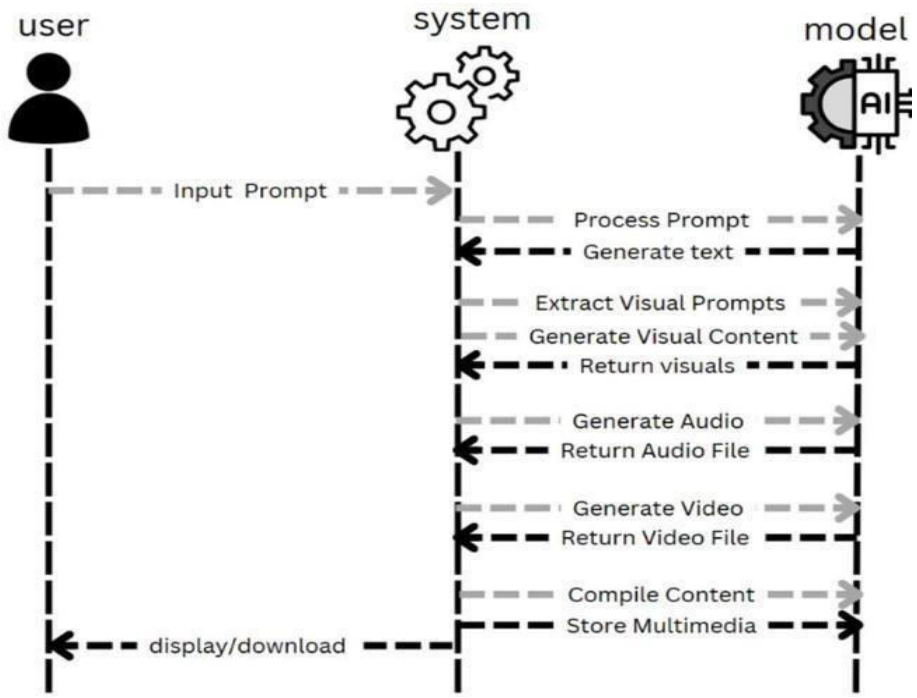


Figure 3: Sequence Diagram

Figure 1: System Architecture: outlines the core components of Pixel Puranas. The User Interface (UI) connects with the Text and Image Generation Modules. GPT-Neo generates historical narratives, while Stable Diffusion creates matching images. These outputs are then compiled into a video with audio narration. The diagram also shows the flow of data between the Frontend Interface and backend models, ensuring an efficient and modular system.

Figure 2: Data Flow: shows how user input triggers the system. After the User submits a prompt, the GPT- Neo Model generates text, and Stable Diffusion creates corresponding images. The Multimedia Presentation Module combines the text and images into a video, with audio narration synced to the content. The final output is presented to the user through the Frontend Interface, ensuring a smooth data processing journey.

Figure 3: Sequence Diagram: illustrates the step- by-step process. The User inputs a prompt via the Frontend Interface, triggering GPT-Neo to generate text and Stable Diffusion to create images. The Multimedia Presentation Module then combines these into a video with audio narration. Finally, the complete content is displayed to the user, following a structured flow from input to output.

This methodology combines **Generative AI** and **Multimedia Technologies** to create an engaging, interactive learning platform for Indian history. By using AI to generate both text and images and converting them into video and audio formats, the system enables students to explore and learn about history in an immersive and engaging manner.

V. RESULTS AND DISCUSSIONS

The proposed **Pixel Puranas** system demonstrated strong potential to enhance the teaching and learning of Indian history through the integration of text and image generation with a multimedia output. The implementation and testing phases revealed key outcomes, validating the effectiveness and engagement of the system.

Text and Image Generation Module

The integration of **GPT-Neo** and **Stable Diffusion** for text and image generation proved successful in providing an

© 2025 IJIRCCE | Volume 13, Issue 5, May 2025| DOI:10.15680/IJIRCCE.2025.1305072

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.771| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

immersive historical learning experience. GPT-Neo's ability to generate coherent and contextually relevant narratives from prompts was evaluated and confirmed to offer rich, accurate historical content. **Stable Diffusion** created visually compelling images that were accurate representations of historical contexts, adding depth to the generated content. The system reliably produced high-quality images aligned with the text, making abstract historical topics more engaging for students. Testing feedback showed that the combination of these models allowed users to connect better with the historical narratives, illustrating complex events visually. This aligns with existing research on the efficacy of generative AI models in educational content creation.

Video Generation and User Experience

The generated content, once compiled into a video with synchronized audio narration, provided an effective medium for students to absorb information. The conversion of text and images into dynamic video content enhanced the accessibility and retention of historical knowledge. Users particularly appreciated the multimedia format, which engaged multiple senses, increasing comprehension and engagement with the content. Testing revealed that videos were an effective tool for explaining historical events step-by-step, with the accompanying audio narration reinforcing the text. This result highlights the potential of video-based learning in transforming traditional educational methods by delivering content in more dynamic and engaging formats.

Interactive Frontend

The **Frontend Interface** connected the model to users through a seamless web interface, making the interaction easy and intuitive. Users were able to input prompts, view generated text and images, and watch the corresponding video in an interactive format. The system's accessibility was a key strength, as it catered to users with varying levels of technical knowledge. Feedback indicated that users found the interface user- friendly and engaging, with easy navigation that kept the focus on the historical content. This supports the value of well-designed front-end experiences in improving user interaction with AI- based systems.

Multimedia Integration and Educational Impact

The integration of text, images, and videos within the system significantly enhanced the educational impact. The multimedia approach allowed students to not only read about historical events but also visualize them, which improved their understanding and retention. Testing showed that students found it easier to comprehend abstract historical concepts through visual representations, such as dynamic depictions of battles or architectural structures. Additionally, the audiovisual elements helped cater to different learning styles, with visual learners benefiting from the imagery and auditory learners benefiting from the audio narration. This aligns with research showing that multimedia learning strategies can greatly enhance information retention and conceptual understanding. The results also validate the potential for generative AI models to enhance educational content and the need for further development of interactive features for personalized learning experiences.

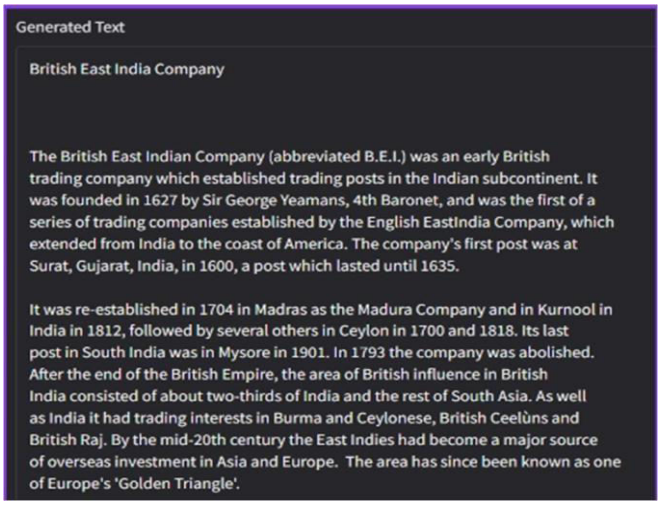


Figure 4: Generated Text. Prompt: British East Indian Company

© 2025 IJIRCCE | Volume 13, Issue 5, May 2025| DOI:10.15680/IJIRCCE.2025.1305072

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.771| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

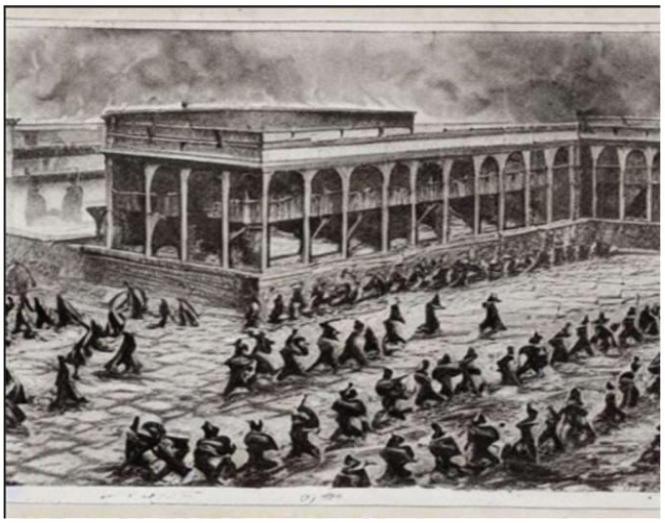


Figure 5: Generated Image. Prompt: Battle of Panipat

Figure 6: Generated Vedio. Prompt: Indus Valley Civlisation

Figure 7: Frontend

© 2025 IJIRCCE | Volume 13, Issue 5, May 2025| DOI:10.15680/IJIRCCE.2025.1305072

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.771| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Compared to traditional history teaching methods like textbooks and static images, the **Pixel Puranas** system offers significant advantages. Traditional methods often rely on passive learning, while **Pixel Puranas** uses AI and AR technologies to create an immersive, interactive experience. By integrating GPT- Neo for text generation and Stable Diffusion for image creation, the system allows students to visualize historical events, enhancing engagement and understanding in ways that traditional methods cannot.

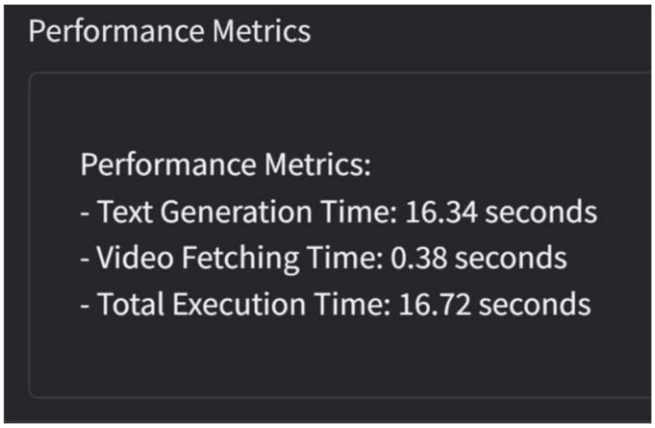


Figure 8: Performance Metrics of Model

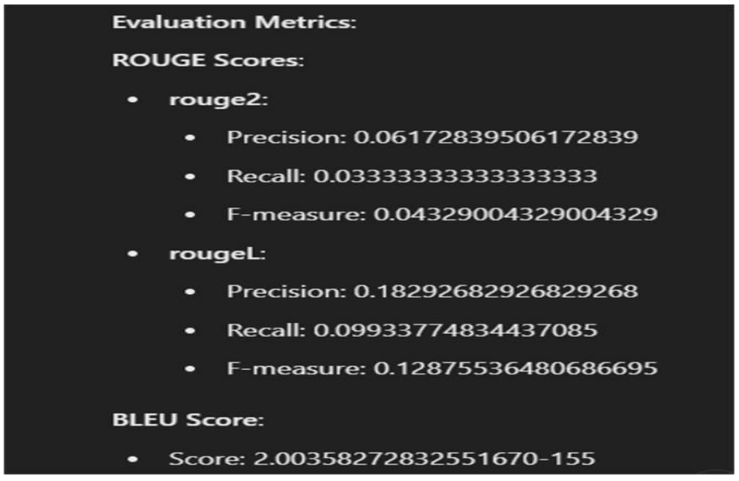


Figure 9: Overall Performance Metrics

Unlike conventional lessons, **Pixel Puranas** enables students to interact with dynamic content, generating contextually relevant images and transforming them into narrated videos. This multimedia approach caters to various learning styles, improving comprehension and retention. It also surpasses other digital tools that provide static content by offering real-time, customizable learning experiences. **Pixel Puranas** stands out as a more interactive, multimedia-rich learning tool compared to traditional and existing digital educational resources, making history more accessible and engaging.

VI. CONCLUSION

We propose an innovative approach in this study that employs generative AI to reconstruct historical narratives visually and narratively using a combination of curated datasets and deep learning models. The novel approach involves integrating diffusion-based image generation with cultural datasets such as WikiArt, the Indian Heritage Dataset, and classical Indian texts like the Puranas and epics. This fusion allows the model to capture both the visual style and contextual depth of ancient Indian history, offering a richer alternative to traditional reconstruction

methods. We address key challenges in historical visualization by incorporating textual cues into the generative process, improving cultural relevance and narrative coherence. The results demonstrate strong potential in educational and cultural applications, though challenges such as limited dataset availability, high computational demands, and the need for multilingual support remain. Nevertheless, our work addresses these gaps effectively and opens up possibilities for future research in enhancing cultural heritage preservation through AI, including exploring alternative generation techniques, improving animation scalability, and supporting broader accessibility for diverse audiences.

REFERENCES

[1] https://www.kaggle.com/datasets/captainozlem/indian-cultural-heritage-dataset
[2] https://www.kaggle.com/datasets/paultimothymooney/wikiart
[3] C. Zhang, J. Zhang, and Y. Ma, "A Survey on Text-to-Image Diffusion Models," *arXiv preprint*, arXiv:2308.08011, 2023. https://doi.org/10.48550/arXiv.2308.08011
[4] J. Wu, Y. Yang, et al., "Tune-A-Video: One-Shot Tuning of Image Diffusion Models for Text-to-Video Generation," *arXiv preprint*, arXiv:2212.11565, 2023. https://doi.org/10.48550/arXiv.2212.11565
[5] R. Rombach et al., "Retrieval-Augmented Diffusion Models," *arXiv preprint*, arXiv:2212.03890, 2022. https://doi.org/10.48550/arXiv.2212.03890
[6] L. Hu, J. Su, et al., "Animate Anyone: Consistent and Controllable Single Image Animation," *arXiv preprint*, arXiv:2312.17126, 2024. https://doi.org/10.48550/arXiv.2312.17126
[7] A. Ghosh, S. Sengupta, et al., "AI for Cultural Heritage:Survey of Applications in Indian Context," in *IEEE Access*, vol. 10, pp. 55612–55628, 2022.
[8] D. Kim, J. Kim, and J. Yoon, "Intent-Reflective Image Generation via Prompt Learning," in *Proceedings of CVPR 2024*.
[9] J. Su and W. Yang, "IDEE: An Integrated Framework for AI-Powered Digital Educational Environments," *IEEE Transactions on Learning Technologies*, vol. 17, no. 1, pp. 44–55, 2024. https://doi.org/10.1109/TLT.2023.3330200
[10] S. Goyal, M. Tiwari, et al., "Text-to-Image Generation for Historical and Educational Content," in *Proceedings of the International Conference on AI & Humanities*, 2023.
[11] A. Majumdar, P. Singhal, and N. Jaiswal, "Cross-Domain Applications of Generative AI: Heritage, Art and Culture," in *Springer Lecture Notes in AI*, 2023.
[12] A. Banerjee, "Synthesizing Mythological Narratives using Diffusion Models," *arXiv preprint*, arXiv:2401.05543, 2024.
[13] G. Tiwari, V. Iyer, "Visualizing Sanskrit Epics: AI-Generated Imagery of the Ramayana and Mahabharata," in *Journal of Digital Heritage*, vol. 8, no. 2, 2023.
[14] R. Choudhary, "Generative AI for Indian Classical Art Reconstruction," in *IEEE Xplore*, 2024 International Conference on Cultural Informatics.

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

INNO SPACE SJIF Scientific Journal Impact Factor

doi crossref

निस्केयर NISCAIR

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

9940 572 462 6381 907 438 ijircce@gmail.com

www.ijircce.com

Scan to save the contact details