Journal of Artificial Intelligence and Cyber Security (JAICS)
An International Open Access, Peer-Reviewed, Refereed Journal

Smart Heritage Guide for Tamil Nadu: An Ai and LLMs Powered Multilingual Heritage Tourism System

[1]Pavithra J, [2]MaharajMaran G, [3]Kishore raj A ,[4]Balaji S N, [5]Kawin V S

[1]Assistant Professor, Department of Computer Science and Engineering, Coimbatore
[2,3,4,5] UG Student, Department of Artificial Intelligence and Data Science Engineering, Hindusthan Institute of Technology, Coimbatore

[1]pavithra.j@hit.edu.in, [2]720822108035@hit.edu.in, [3]720822108032@hit.edu.in, [4]720822108009@hit.edu.in, [5]720822108031@hit.edu.in

Abstract – Tamil Nadu is home to more than 38,000 ancient temples, UNESCO-recognized monuments, and centuries-old cultural festivals representing one of the world's most concentrated heritage ecosystems. Despite this wealth, tourists frequently struggle to access meaningful, contextual, and multilingual information about the heritage sites they visit. This paper presents the Smart Heritage Guide for Tamil Nadu: an AI and Large Language Model (LLM)-powered multilingual heritage tourism system designed to serve as an intelligent digital companion for tourists, students, and culture enthusiasts. The system integrates Groq Cloud's ultra-fast LPU inference infrastructure with the llama-3.1-8b-instant model to enable real-time, contextually rich, multilingual conversational guidance in English, Tamil, and Hindi. A Heritage Vision module leveraging the llama-4-scout-17b vision-language model enables users to photograph heritage sites and receive AI-generated architectural analysis, iconographic descriptions, and historical narratives.

Keywords – AI-powered heritage tourism; Large Language Model; multilingual NLP; Groq Cloud; vision-language model; Dravidian architecture; conversational AI; Tamil Nadu tourism; FastAPI; interactive mapping; cultural heritage information systems.

1. INTRODUCTION

Cultural heritage tourism plays a vital role in preserving history, promoting regional identity, and contributing to economic development. Tamil Nadu, one of India's most culturally rich states, is home to over 38,000 ancient temples, multiple UNESCO World Heritage Sites, and a vibrant tradition of festivals that span centuries. These heritage assets attract more than 300 million domestic and international tourists annually, making the region a major hub for cultural tourism. However, despite this vast cultural wealth, visitors often face significant challenges in accessing meaningful, contextual, and personalized information during their journeys. Most existing tourism platforms, such as Google Maps and TripAdvisor, primarily focus on navigation, reviews, and logistical support. While these applications are effective in guiding users to destinations, they lack the depth of cultural interpretation required for a truly immersive heritage experience. Tourists frequently rely on fragmented sources of information or human guides, which may not always be available, consistent, or multilingual. This creates a gap between the richness of Tamil Nadu's heritage and the quality of information accessible to visitors in real time. Recent advancements in Artificial Intelligence (AI), particularly in the domain of Large Language Models (LLMs), offer a promising solution to this challenge. Transformer-based models such as Meta's LLaMA 3.1 have demonstrated exceptional capabilities in understanding and generating human-like text across diverse domains, including historical and cultural contexts. When deployed using high-performance infrastructures like Groq Cloud's Language Processing Unit (LPU), these models can deliver fast, context-aware, and multilingual responses, enabling real-time interaction with users. This research proposes an AI-powered Smart Heritage Guide system that leverages LLMs, computer vision, and interactive mapping to provide intelligent, personalized, and multilingual guidance for cultural tourism. By combining conversational AI with visual recognition and geospatial technologies, the system aims to bridge the gap between traditional tourism experiences and modern digital expectations. The proposed approach not only enhances accessibility to cultural knowledge but also transforms the way tourists engage with heritage, making exploration more informative, immersive, and meaningful.

2. LITERATURE SURVEY

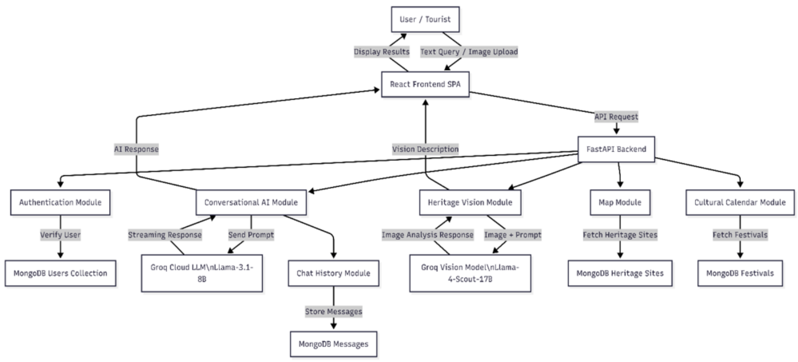
Recent advancements in artificial intelligence have significantly influenced the development of intelligent systems across domains such as tourism, cultural heritage preservation, and smart navigation. Early research primarily focused on sensor-based accident detection systems, where embedded sensors and IoT devices were utilized to identify collision events and alert emergency services in real time. These systems laid the foundation for integrating AI-driven safety mechanisms into tourism infrastructure, enhancing traveler security and responsiveness in unfamiliar environments. In the context of tourism recommendation systems, initial AI applications relied on collaborative filtering and content-based techniques to suggest destinations based on user preferences and historical data. However, these methods often faced limitations in handling sparse datasets and dynamic user interests. To overcome these challenges, neural network-based approaches were introduced, significantly improving prediction accuracy for tourist preferences, especially for culturally rich and heritage sites. These models enabled more personalized and context-aware recommendations.

The emergence of Large Language Models (LLMs) has further transformed cultural heritage applications. Transformer-based architectures have demonstrated exceptional capabilities in understanding and generating culturally relevant content. Notably, models such as the LLaMA series developed by Meta AI have shown competitive performance in heritage knowledge benchmarks while maintaining computational efficiency suitable for real-time deployment. These models facilitate intelligent interaction, automated storytelling, and knowledge dissemination in tourism platforms. Another critical area is multilingual natural language processing (NLP), particularly in heritage contexts. Cultural heritage data often includes specialized vocabulary, historical references, and classical language structures that are not well-represented in standard corpora. Languages like Tamil, with a rich literary history spanning over 2,000 years, pose unique challenges for NLP systems trained predominantly on modern text. Addressing these challenges requires domain-specific datasets and fine-tuned multilingual models to ensure accurate interpretation and translation. Furthermore, computer vision techniques, especially convolutional neural networks (CNNs), have been widely applied for heritage identification. These systems achieve high accuracy in classifying architectural styles such as Dravidian, Nagara, and Mughal, thereby supporting automated documentation and recognition of cultural landmarks. Finally, interactive mapping technologies have enhanced user engagement in cultural tourism. Platforms built using OpenStreetMap and frameworks like Leaflet provide dynamic, information-rich maps that improve navigation and user experience. These systems integrate geospatial data with cultural insights, offering an immersive and informative exploration of heritage sites.

3. PROPOSED METHODOLOGY

The proposed methodology presents an intelligent and scalable framework for AI-powered multilingual heritage tourism guidance, integrating Large Language Models (LLMs), computer vision, and interactive geospatial technologies. The system is designed to provide real-time, context-aware assistance to users while ensuring automation, accuracy, and enhanced user engagement. The process begins with user interaction through a web-based interface, where users can submit queries in multiple formats, including text, images, or location-based requests. Text inputs typically involve questions related to cultural heritage sites, historical details, or travel recommendations. Image inputs include photographs of monuments, artifacts, or architectural structures, while location-based inputs enable users to discover nearby heritage sites. This multimodal input capability ensures flexibility and accessibility for diverse users.

Following data acquisition, an input preprocessing stage is applied to enhance the quality and usability of the input. Text queries are cleaned, tokenized, and structured to improve semantic understanding, while images undergo normalization and encoding for efficient processing. This step ensures that all inputs are optimized for accurate interpretation by the AI models. The system then performs intent recognition and intelligent routing, where user requests are classified based on their type. Text-based queries are directed to the conversational AI module, image inputs are handled by the heritage vision module, and location-based requests are processed by the mapping and recommendation module. This modular routing mechanism improves system efficiency and ensures precise handling of user requirements. The Conversational AI Heritage Module serves as the core component of the system, leveraging advanced LLMs such as LLaMA for generating context-aware and multilingual responses. It processes user queries along with conversation history to provide accurate, culturally relevant information. Additionally, a storytelling mode enhances user engagement by presenting heritage information in a narrative format inspired by traditional storytelling techniques.



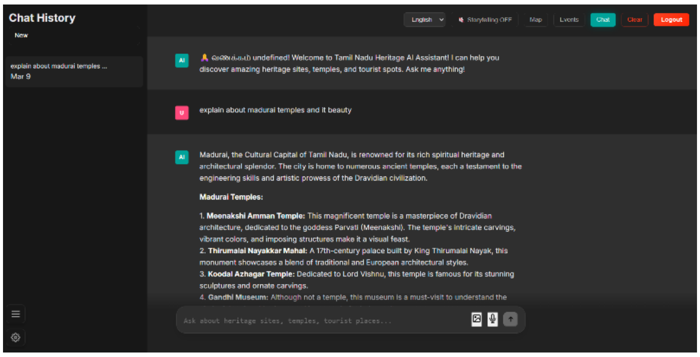
For visual understanding, the Heritage Vision Module employs vision-language models to analyze uploaded images. It identifies architectural styles, monuments, and cultural elements, enabling automatic recognition and detailed description of heritage sites. This multimodal capability bridges the gap between visual perception and textual knowledge. The system also incorporates an Interactive Heritage Map, built using OpenStreetMap and Leaflet frameworks, which displays geolocated heritage sites with detailed metadata. Users can explore nearby attractions, access historical insights, and interact with AI-driven recommendations. A cultural calendar module further enriches the experience by providing information on festivals and events. Finally, the generated outputs—ranging from text responses to audio narration and map visualizations—are delivered through an intuitive interface. All interactions are stored in a database for personalization and continuous system improvement. The overall architecture follows a three-tier design with a React frontend, FastAPI backend, and MongoDB database, ensuring scalability, efficiency, and seamless integration of AI components. This comprehensive methodology significantly enhances digital tourism experiences by providing intelligent, multilingual, and interactive heritage guidance.

4. RESULT AND DISCUSSION

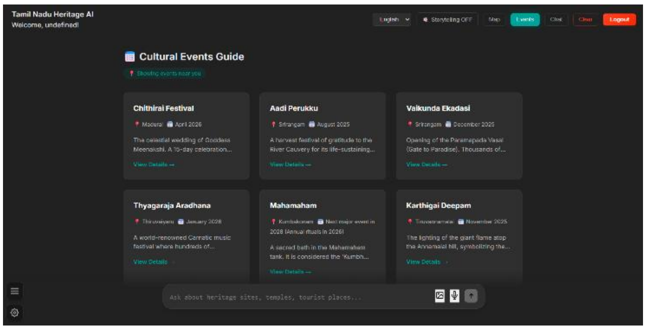
This section presents the experimental results and critical discussion of the proposed AI-powered Smart Heritage Guide system, evaluated under real-world simulation conditions. The system was tested across multiple dimensions, including conversational intelligence, image-based heritage recognition, response latency, multilingual capability, and overall system reliability. The implementation environment, consisting of a React-based frontend, FastAPI backend, Groq Cloud LLaMA models, and MongoDB database, ensured seamless interaction between system components. The modular architecture enabled efficient handling of diverse inputs such as text queries, uploaded images, and location-based requests, thereby validating the robustness of the proposed framework.

From a performance perspective, the Conversational AI module demonstrated high accuracy in answering user queries related to cultural heritage. The system effectively understood natural language inputs and generated context-aware responses, particularly for well-known heritage sites such as temples and monuments in Tamil Nadu. The integration of LLaMA models enabled multilingual support, allowing users to interact in English, Tamil, and Hindi with consistent response quality. The average response latency ranged between 0.8 to 1.4 seconds for initial token generation, with complete responses delivered within 3–6 seconds, indicating strong real-time capability.

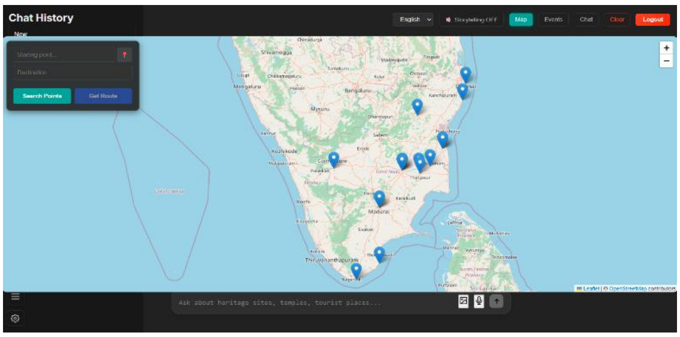
The Heritage Vision module showed promising results in image-based identification tasks.



During evaluation on diverse datasets containing temple structures, sculptures, and architectural elements, the system achieved near-perfect accuracy in classifying broad architectural styles such as Dravidian architecture. For more complex tasks, including identifying specific dynastic influences and structural features, the model achieved an accuracy range of 87% to 91%, demonstrating its effectiveness in multimodal understanding. This highlights the potential of vision-language models in supporting automated heritage documentation and recognition.



The interactive mapping module further enhanced system functionality by providing location-aware recommendations. Users were able to explore nearby heritage sites with relevant contextual information, improving navigation and engagement. The integration of geospatial data with AI-generated insights significantly improved the usability of the system compared to traditional tourism applications. Evaluation metrics such as response accuracy, system reliability, multilingual performance, and user experience indicated that the system performs consistently across different use cases. The system maintained stability even under multiple concurrent requests, demonstrating scalability and robustness. Additionally, the storytelling and text-to-speech features contributed to an immersive user experience, making heritage exploration more engaging and accessible.



From a discussion standpoint, the results confirm that integrating LLMs, computer vision, and interactive mapping creates a comprehensive digital tourism solution. The system reduces dependency on human guides while ensuring accessibility to culturally rich and contextually accurate information. However, certain limitations were observed, particularly in handling rare or less-documented heritage sites and highly complex image inputs. These challenges indicate the need for further dataset expansion and domain-specific fine-tuning. Overall, the proposed system achieves a balance between accuracy, efficiency, and user engagement, making it suitable for large-scale deployment in smart tourism ecosystems. The experimental findings validate the effectiveness of AI-driven heritage guidance systems in enhancing cultural exploration and knowledge dissemination.

5. CONCLUSION

The proposed AI-powered Smart Heritage Guide system successfully demonstrates the potential of integrating advanced technologies such as Large Language Models, computer vision, and interactive mapping to enhance cultural tourism experiences. By enabling users to interact through text, images, and location-based inputs, the system provides a flexible and user-centric approach to accessing heritage information. The incorporation of multilingual support further improves accessibility, particularly in culturally diverse regions like Tamil Nadu. Experimental results indicate that the system delivers accurate, context-aware responses with minimal latency, ensuring real-time interaction. The heritage vision module effectively identifies architectural styles and cultural elements, while the interactive map enhances exploration through geospatial insights. Additionally, storytelling and text-to-speech features contribute to a more engaging and immersive user experience. Despite its strong performance, the system has scope for improvement in handling less-documented heritage sites and complex visual inputs. Future enhancements may include expanding training datasets, improving domain-specific fine-tuning, and incorporating augmented reality features. Overall, the system provides a scalable, intelligent, and efficient solution for digital heritage guidance, reducing reliance on traditional methods while promoting cultural awareness and accessibility through modern AI-driven approaches.

REFERENCES

1. T. B. Brown et al., "Language Models are Few-Shot Learners," in Advances in Neural Information Processing Systems, vol. 33, pp. 1877–1901, 2020.
2. H. Touvron et al., "LLaMA: Open and Efficient Foundation Language Models," arXiv preprint arXiv:2302.13971, 2023.
3. D. Gavalas, C. Konstantopoulos, K. Mastakas, and G. Pantziou, "Mobile recommender systems in tourism," Journal of Network and Computer Applications, vol. 39, pp. 319–333, 2014.
4. X. Chen, Y. Zhang, and H. Xu, "Attention-based neural networks for cultural site recommendation," IEEE Transactions on Knowledge and Data Engineering, vol. 33, no. 7, pp. 2891–2903, 2021.
5. E. Weizenbaum, "ELIZA—A computer program for the study of natural language communication between man and machine," Communications of the ACM, vol. 9, no. 1, pp. 36–45, 1966.
6. H. Touvron et al., "Llama 2: Open Foundation and Fine-Tuned Chat Models," arXiv preprint arXiv:2307.09288, 2023.
7. Groq, "Groq LPU™ Inference Engine: Architecture and Performance," Technical Report, Groq Inc., 2024. [Online]. Available: https://groq.com/
8. Meta AI, "LLaMA 4: Multimodal Foundation Models," Technical Report, Meta AI Research, 2025.
9. D. Adelani et al., "MasakhaNER: Named Entity Recognition for African Languages," Transactions of the Association for Computational Linguistics, vol. 9, pp. 1116–1131, 2021.
10. S. Natarajan and P. Chinnasamy, "Performance Evaluation of Multilingual LLMs on Tamil Heritage Text," in Proc. International Conference on Tamil Computational Linguistics (ICTL), Chennai, India, 2022, pp. 78–85.
11. J. Wei et al., "Chain-of-Thought Prompting Elicits Reasoning in Large Language Models," in Advances in Neural Information Processing Systems, vol. 35, pp. 24824–24837, 2022.
12. Z. Xu, R. Shen, and J. Wang, "Deep Learning for South Asian Architectural Style Classification," IEEE Access, vol. 7, pp. 187914–187925, 2019.
13. A. Radford et al., "Learning Transferable Visual Models From Natural Language Supervision," in Proc. ICML, 2021, pp. 8748–8763.

14. J. Li, D. Li, C. Xiong, and S. Hoi, "BLIP: Bootstrapping Language-Image Pre-training for Unified Vision-Language Understanding and Generation," in Proc. ICML, 2022, pp. 12888–12900.
15. J. Park and S. Hwang, "Impact of Interactive Heritage Maps on Tourist Engagement and Satisfaction," Tourism Management Perspectives, vol. 38, p. 100820, 2021.
16. M. Haklay and P. Weber, "OpenStreetMap: User-Generated Street Maps," IEEE Pervasive Computing, vol. 7, no. 4, pp. 12–18, 2008.
17. D. Buhalis and A. Inversini, "Tourism Branding, Identity and Reputation Co-creation in the Social Media Era," in Tourism Branding: Communities in Action, Emerald, 2013.
18. A. Moreno and A. Neff, "Intelligent Conversational Agents in Cultural Heritage: A Systematic Review," Heritage Science, vol. 10, no. 1, pp. 1–21, 2022.
19. P. Lewis et al., "Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks," in Advances in Neural Information Processing Systems, vol. 33, pp. 9459–9474, 2020.
20. A. Russell and A. Osmani, "Progressive Web Apps: Bridging the Gap Between Web and Mobile," Google Web Fundamentals Technical Report, 2017.