

Personalized Generative Storytelling with AI-Visual Illustrations for the Promotion of Knowledge in Cultural Heritage Tourism

Andrea Ferracani ♠ Marco Bertini ♠ Pietro Pala ♠ Gabriele Nannotti ♠ Filippo Principi ♠
Giuseppe Becchi ♠
♠{firstname.lastname}@unifi.it

Introduction

This poster presents a **mobile application** designed to enhance cultural tourism through interactive storytelling and AI-generated images. Using GPT-4 and a **custom image generation pipeline**, the app allows tourists to create personalized city tours, where they become the protagonist of a narrative. Users can select points of interest (POIs) and narrative genres, with real-time visual illustrations enhancing the immersive experience. Technical challenges addressed include generating coherent stories and high-quality images in real-time. **User studies** demonstrate the educational and experiential benefits of this AI-driven approach to cultural heritage tourism.

Contributions

- Interactive digital storytelling:** we present a mobile application that leverages GPT-4 to create personalized, AI-driven narratives for cultural tourism, allowing users to choose genres and Points of Interest (POIs) while being the protagonist of their own story.
- Generative visual illustrations:** the application integrates a custom image generation pipeline that produces visual representations of user actions within the story, ensuring coherence and enhancing immersion through high-quality, real-time images.
- Technical advancements in storytelling:** our system addresses challenges such as producing coherent narratives, real-time image generation, and maintaining identity consistency, using advanced prompt engineering and Low-Rank Adaptation (LoRA) models.
- Personalization and immersion:** the app allows users to personalize their experience by defining character appearances and story settings, offering a dynamic and highly immersive cultural tour experience.
- User study validation:** we validate the app's effectiveness through user studies, which demonstrate the system's capacity to enhance engagement, immersion and educational outcomes in cultural tourism.

Application Overview

The mobile application enhances **cultural tourism** through interactive storytelling and AI-generated images. Users can select **Points of Interest (POIs)**, **story genres**, type of narration and selfies to create personalized city tours where they are the protagonist. The app dynamically generates narratives and visual illustrations based on user input, exploiting selfies for character representation. This provides an immersive, tailored experience, enriching cultural exploration with multimedia content like descriptions, images, panoramic views and videos tied to each POI. Furthermore, **users can control story progression** through automatically generated story options.

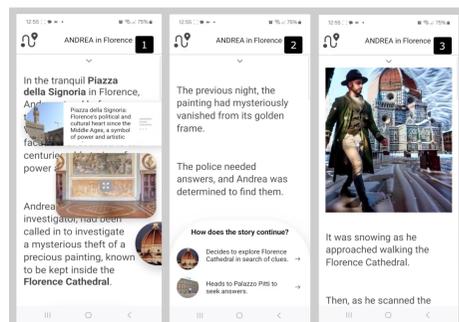


Figure 1: The app: 1) Generated story with contextual POIs; 2) Story progression choices; 3) An example of a generated image contextual to the narrative.

Custom Image Generation Pipeline

The app features a **multi-step AI-powered pipeline** for generating personalized images. Starting with a base image of a location, the pipeline modifies the theme, weather and time of day. It then inserts the user's selfie into the scene, adjusting pose and identity details to maintain consistency. The final step enhances photorealism and integrates the user's face, creating a seamless, immersive visual representation that aligns with the story's narrative.

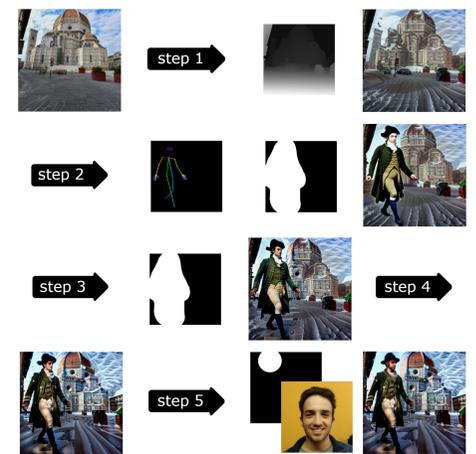


Figure 2: Custom Image Generation Pipeline

The image generation pipeline leverages **state-of-the-art deep generative models**, including **Stable Diffusion** and **ControlNet**, for tasks such as image inpainting, photorealistic enhancement, and the preservation of user identity. These models are optimized to maintain visual coherence and narrative alignment, ensuring high-quality, personalized images in real time.

Evaluation Results

We conducted a thorough evaluation of the app's performance, focusing on user satisfaction, engagement, immersion and learning outcomes.

User Satisfaction, Engagement and Immersion: participants were divided into three groups: G1 (Printed materials), G2 (App Non-Interactive), and G3 (App Interactive). Table 1 shows the comparison of satisfaction, engagement, and immersion across the groups on a five points LikertS scale.

Measure	G1	G2	G3
Satisfaction	6.4	5.5	5.8
Engagement	3.4	5.5	8.3
Immersion	2.9	6.5	7.8

Table 1: Comparison of satisfaction, engagement and immersion between groups

The interactive version of the app (G3) achieved the highest scores in both engagement and immersion, highlighting the benefits of interactivity and personalized images in enhancing user experience.

Learning Outcomes Assessment: we assessed participants' learning outcomes using a multiple-choice questionnaire consisting of **10 questions** about the cultural heritage of the Points of Interest (POIs) presented in the app's stories. The results showed that **G3** achieved the highest average score, with **7.8 correct answers out of 10**. In comparison, **G1** scored an average of **5.8**, while **G2** scored **4.5**. These results suggest that the interactive app, which includes personalized AI-generated stories and images, significantly improves the retention of cultural information compared to traditional printed materials or a non-interactive version of the app. The interactive features and personalized visual content appear to enhance user engagement and learning.

Acknowledgements