VIRTUAL 3D TOUR

Project Team

Dabeer Ul Haq Qureshi	21I-1518
Senan Faisal	21I-1546
Saad Ahsan	21I-2965

Session 2021-2025

Supervised by

Mr Zeshan Khan

Co-Supervised by

Mr./ Ms./ Dr. Supervisor Name

Department of Computer Science

National University of Computer and Emerging Sciences
Islamabad, Pakistan

June, 2025

Contents

List of Figures

List of Tables

Chapter 1

Introduction

The purpose of this project proposal document is to outline the objectives and approach for the "Virtual 3D Tour" project, which aims to enhance tourism in Pakistan's northern areas by converting 2D images into immersive 3D visualizations. The project leverages advanced technologies like Convolutional Neural Networks (CNN), Large Language Models (LLM), and Neural Radiance Fields (NeRF) to create photorealistic 3D models from 2D images, providing users with an interactive and realistic virtual tour experience through a mobile app. This project seeks to boost tourism, promote the region's beauty, and set a new standard for virtual tourism.[1].

1.1 Problem Statement

The Virtual 3D Tour project aims to address the challenge of limited accessibility to Pakistan's northern areas due to geographical constraints and lack of awareness about their natural beauty and cultural significance. By converting 2D images into immersive 3D visualizations, this project seeks to provide a virtual tour experience that allows users to explore these regions from the comfort of their homes, overcoming geographical barriers and fostering a deeper appreciation for their unique attractions. This virtual platform will not only promote tourism but also serve as an educational tool, showcasing the diverse landscapes, historical sites, and local traditions of Pakistan's northern areas to a wider audience.

1.2 Motivation

The primary motivation for developing the Virtual 3D Tour project lies in the untapped potential of Pakistan's northern areas as tourist destinations. Despite their breathtaking

beauty and rich cultural heritage, these regions often remain under-explored due to several factors:

Geographical Barriers: The mountainous terrain and remote locations of these areas can make them difficult to access for many travelers. Lack of Awareness: Many people are unaware of the diverse attractions and unique experiences that the northern areas offer. Limited Infrastructure: The region's infrastructure, including transportation and accommodation, may not be as developed as other tourist destinations. By creating a virtual tour experience, we aim to address these challenges and make these regions more accessible to a wider audience. This will not only boost tourism but also contribute to the economic development and cultural preservation of the northern areas.

1.3 Problem Solution

Objectives:

Enhance Accessibility: Provide a virtual tour experience that allows users to explore Pakistan's northern areas from the comfort of their homes, overcoming geographical constraints. Promote Tourism: Showcase the natural beauty, cultural heritage, and unique attractions of the northern areas to a wider audience, stimulating tourism. Preserve Cultural Heritage: Contribute to the preservation of local traditions and customs by promoting awareness and appreciation for the region's cultural diversity. Educational Tool: Serve as a valuable educational resource, providing information about the region's history, geography, and ecology. Economic Development: Boost the local economy by attracting more tourists and supporting related businesses. Goals:

Develop a user-friendly and immersive virtual tour experience that accurately represents the landscapes, landmarks, and cultural features of the northern areas. Create a platform that is easily accessible to users with different technological backgrounds and devices. Generate interest and excitement about visiting the northern areas in person. Foster a sense of connection and appreciation for the region's natural beauty and cultural heritage among users. Provide a sustainable and scalable solution for promoting tourism in the northern areas. Application of Software:

The Virtual 3D Tour software will leverage advanced technologies such as Convolutional Neural Networks (CNN), Large Language Models (LLM), and Neural Radiance Fields (NeRF) to achieve these objectives and goals. Specifically, the software will:

Process and Analyze Images: Use CNNs to extract relevant features from 2D images of the northern areas. Generate 3D Models: Employ NeRF to create photorealistic 3D models based on the extracted features, providing users with a realistic and immersive experience. Provide Informative Content: Utilize LLMs to generate informative descrip-

tions and captions for the 3D models, enhancing the user's understanding of the region's history, culture, and attractions. Develop a User-Friendly Interface: Create a mobile app or web-based platform with intuitive navigation and interactive features, allowing users to explore the 3D models seamlessly. By effectively applying these technologies, the Virtual 3D Tour software will provide a comprehensive and engaging experience that meets the needs and expectations of users, while also achieving the broader goals of promoting tourism and preserving cultural heritage in Pakistan's northern areas.

(Usually in 14-16 sentences)

1.4 Stake Holders

The Virtual 3D Tour project involves a diverse range of stakeholders, each with distinct interests and perspectives. These key stakeholders include:

Government Agencies: Tourism boards, local administrations, and cultural heritage departments in the northern areas. Tourism Industry: Travel agencies, tour operators, hotels, and restaurants operating in the region. Local Communities: Residents of the northern areas who benefit from tourism-related activities and economic development. Cultural Experts: Historians, archaeologists, and anthropologists with knowledge of the region's cultural heritage. Technology Providers: Companies specializing in 3D modeling, virtual reality, and mobile app development. Potential Users: Tourists, both domestic and international, who are interested in exploring Pakistan's northern areas. Academic Institutions: Universities and research organizations interested in the application of technology in tourism and cultural preservation. Funding Agencies: Organizations that provide financial support for the project, such as government grants, private foundations, or corporate sponsors. By engaging with these stakeholders throughout the project development process, we can ensure that the Virtual 3D Tour aligns with their needs and contributes to the overall goals of promoting tourism and preserving cultural heritage in Pakistan's northern areas.

Chapter 2

Project Description

2.1 Scope

The Virtual 3D Tour project aims to create an immersive and interactive experience for users to explore Pakistan's northern areas. The scope of the project encompasses the following key functionalities:

Data Collection and Processing: Gathering a comprehensive dataset of high-quality 2D images from various tourist destinations in the northern areas, followed by preprocessing and cleaning the images to ensure optimal quality for 3D model generation.

3D Model Generation: Utilizing advanced technologies like Convolutional Neural Networks (CNN) and Neural Radiance Fields (NeRF) to reconstruct photorealistic 3D models from the collected images, capturing the intricate details and unique characteristics of each destination.

Mobile App Development: Creating a user-friendly mobile app that allows users to seamlessly navigate and interact with the generated 3D models. The app will incorporate features such as zooming, panning, and rotation, providing a realistic and immersive virtual tour experience.

Informative Content Integration: Incorporating informative descriptions, captions, and historical context for each destination, generated using Large Language Models (LLM), to enhance the user's understanding and appreciation of the region's cultural heritage and natural beauty.

User Interface Design: Designing an intuitive and visually appealing user interface that is easy to navigate and provides a seamless user experience.

Testing and Refinement: Thoroughly testing the mobile app to ensure its functionality, performance, and compatibility with various devices. Gathering user feedback to identify areas for improvement and make necessary refinements to enhance the overall user

experience.

Scalability: Designing the system to accommodate future expansion and updates, allowing for the addition of new destinations and features as the project evolves.

By focusing on these key functionalities, the Virtual 3D Tour project will provide a valuable and engaging tool for promoting tourism in Pakistan's northern areas, showcasing their natural beauty and cultural heritage to a wider audience.

2.2 Modules

Write down the modules of the proposed project. Each module should highlight features, using bulleted/numbered notation. When developing both a mobile app and a web app, group the modules according to the system types, such as, Client Web App, Client Mobile App, Admin Web App etc.

2.2.1 Module 1

Data Acquisition and Processing

1. Collect a diverse dataset of high-quality 2D images from various tourist destinations in Pakistan's northern areas.
2. Clean and preprocess the images to ensure optimal quality for 3D model generation.

2.2.2 Module 2

3D Model Generation

1. Utilize advanced technologies to create photorealistic 3D models from the processed images.
2. Capture the intricate details and unique characteristics of each destination.

2.2.3 Module 3

Mobile App Development

1. Design an intuitive and visually appealing user interface for the mobile app.
2. Incorporate interactive features to provide a realistic and engaging virtual tour experience.

2.2.4 Module 4

Content Integration

1. Generate informative descriptions and captions for each destination using Large Language Models (LLM).
2. Include relevant historical and cultural information to enrich the user's understanding.

2.2.5 Module 5

Testing and Refinement

1. Conduct thorough testing to identify and address any bugs or issues.
2. Gather user feedback to improve the app's performance and user experience.

2.3 Tools and Technologies

- **Hardware Requirements**
- High-performance computing system: A powerful computer with a high-end graphics processing unit (GPU) is essential for training and running the deep learning models used in 3D model generation.
- Large storage capacity: A significant amount of storage space is required to store the dataset, intermediate results, and final 3D models.
- **Software Requirements**
- Deep Learning Framework: TensorFlow or PyTorch are popular choices for implementing Convolutional Neural Networks (CNN) and Neural Radiance Fields (NeRF).
- Image Processing Libraries: OpenCV or PIL for image preprocessing and manipulation.
- 3D Modeling Software: Blender or Unity for creating and visualizing 3D scenes.
- Mobile App Development Tools: Android Studio or Xcode for developing the mobile app.

- Version Control System: Git for managing code changes and collaboration.
- **Cloud Platforms**
- Cloud Computing Services: Consider using cloud platforms like AWS, Google Cloud Platform, or Azure to leverage their scalable computing resources, storage, and machine learning tools.
- **Additional Tools**
- Annotation Tools: LabelImg or VGG Image Annotator for annotating images if required.
- Data Visualization Tools: Matplotlib or Seaborn for visualizing data and results.
- Deployment Tools: Heroku or AWS Amplify for deploying the mobile app.

2.4 Work Division

For each module and respective feature, assign responsibility to a team member

Table 2.1: Table 1: Work Division

Name	Registration	Responsibility / Module / Feature
Dabeer Ul Haq Qureshi	21I-1518	(Module 1 - Feat 1-2) Data Acquisition and Processing (Module 3 - Feat 2) Database tasks
Senan Faisal	21I-1546	(Module 2 - Feat 1-2) 3D Model Generation
Saad Ahsan	21I-2965	(Module 3 - Feat 1) Mobile App Development (Module 5 - Feat 1-2) Testing and Refinement

2.5 TimeLine

Create a timeline relating each iteration to the tasks/modules created previously.

2.6 High-Level Goals

The "Virtual 3D Tour" project aims to achieve the following high-level goals:

1. **Enhance Tourism Experience:** Provide an immersive and interactive virtual experience of tourism destinations in Pakistan's northern areas by transforming 2D images into detailed 3D visuals.

Table 2.2: Table 2

Iteration#	Time Frame	Tasks/Modules
01	Sept	Dataset Collection & Preprocessing, CNN Model Development, Embedding Extraction & Validation
02	Oct-Nov	Advanced CNN Training & Optimization, LLM Integration & Prompt Generation, Mobile App Development Start
03	Dec	3D Model Generations
04	Jan	Mobile App Development & Enhancement, User Testing & Feedback Iteration, Final App Adjustments & Review
05	Feb-Mar	Final Enhancements & Bug Fixes, App Deployment & Final Testing, Project Documentation & Submission

Table 2.3: Comparison of 2D-to-3D Conversion Research with Our Project

Aspect	Our Project	Wu et al. (2017)	Wang et al. (2018)	Guo et al. (2019)
Input Type	2D image of a tourism place	2D images and 3D shapes	Single 2D images	Single 2D images
Intermediate Process	CNN-based embeddings, LLM for prompt generation, 3D visualization	Direct 3D shape generation using GANs	Depth map prediction from 2D images	Depth estimation from single images
Final Output	3D visuals with tourism emphasis	3D object shapes	3D shapes from multi-view depth	3D mesh models
Application and Focus	Tourism sector, highlighting tourist attractions	General object shape generation, computer graphics	Virtual and augmented reality	3D modeling, CAD applications
Technology Stack	Image processing, NLP, 3D rendering	Generative Adversarial Networks (GANs)	Deep generative networks	Deep learning for depth estimation
Similarity	40-50% similarity in 3D visualization goals	-	-	-
Major Differences	Domain-specific focus on tourism, unique NLP integration	General-purpose 3D model generation	Multi-view depth prediction approach	Focus on single image depth estimation

2. **Leverage Advanced Technologies:** Utilize cutting-edge technologies, including Convolutional Neural Networks (CNN) for image analysis, Large Language Models (LLM) for descriptive text generation, and Neural Radiance Fields (NeRF) for

creating photorealistic 3D models.

3. **Develop a User-Friendly Mobile Application:** Create a mobile app that allows users to easily navigate, interact with, and explore 3D models of tourist locations, incorporating features like zooming, panning, and rotation.

4. **Integrate Informative Content:** Enhance the virtual tour experience by integrating informative descriptions, captions, and historical context generated using LLMs to enrich users' understanding of the destinations.

5. **Ensure Robust Testing and Refinement:** Conduct thorough testing of the mobile app to ensure its functionality, performance, and compatibility across various devices. Gather user feedback for continuous improvement and refinement of the app.

6. **Support Future Expansion:** Design the system with scalability in mind, allowing for the addition of new destinations and features as the project evolves.

2.7 Comparison with Terraverse Project

2.7.1 Input Source and Process

Our Project:

- **Input:** A 2D image of a tourism place provided by the user.
- **Process:**
    1. The image is analyzed using a CNN model to generate embeddings, which are detailed descriptions of objects in the image.
    2. These embeddings are used to generate prompts focusing on tourism-related objects.
    3. The prompts guide the creation of 3D visuals of objects related to tourism.

Terraverse Project:

- **Input:** A text prompt or location provided by the user in real-time.
- **Process:**
    1. The text prompt or location is directly converted into 3D visuals using NeRF (Neural Radiance Fields) selection.
    2. NeRF synthesizes 3D scenes from 2D images or textual descriptions.

2.7.2 Core Technologies Used

Our Project:

- **CNN Model:** Used to analyze images and generate embeddings.
- **LLM (Large Language Model):** Used to generate descriptive prompts from embeddings.
- **3D Visuals:** Generated based on tourism-related objects identified in the prompts.

Terraverse Project:

- **Real-time Input Handling:** Takes real-time text prompts or locations.
- **NeRF Selection:** Used to generate 3D visuals directly from textual descriptions or specified locations.

2.7.3 Focus and Application

Our Project:

- Focuses on analyzing visual content (images) and extracting meaningful information to create tourism-focused 3D visuals.
- Emphasizes the step-by-step transformation from 2D images to 3D visuals through intermediate text-based descriptions.

Terraverse Project:

- Focuses on real-time generation of 3D visuals directly from text or location data.
- Skips the image analysis step, going directly from text to 3D visual synthesis.

2.7.4 Similarity in Work

Shared Components:

- Both projects involve generating 3D visuals as an output.
- Both utilize some form of textual description or prompt to guide the 3D visualization.

2.7.5 Major Differences

- **Input Type:** Our project uses 2D images, while the Terraverse project uses text prompts or locations.
- **Process Flow:** Our project has an additional layer of image analysis and intermediate prompt generation, whereas the Terraverse project directly converts text to 3D visuals.
- **Real-time Capability:** The Terraverse project is designed for real-time input, while our project processes images that may not necessarily be provided in real-time.

2.7.6 Estimated Similarity Percentage

- Given the similarities in the final goal (3D visual generation) but differences in input, processing, and real-time capabilities, approximately 40-50% of the work might be considered similar. This includes the conceptual overlap in generating 3D visuals based on extracted or provided descriptions but differs significantly in the methods and input sources.

2.7.7 Conclusion

- While both projects share the common objective of generating 3D visuals, the key differentiators are in how they process inputs and the types of inputs they handle.
- Our project focuses on transforming visual data (images) into 3D outputs via descriptive prompts, while the Terraverse project emphasizes direct text-to-3D conversion in real-time.

2.8 Conclusions

In this work, we presented the "Virtual 3D Tour" project, which focuses on transforming 2D images of tourism destinations into interactive 3D visuals. The key findings and conclusions from our project are as follows:

- **Innovative Approach:** Our project introduces a novel approach by combining Convolutional Neural Networks (CNN) for image analysis with Large Language Models (LLMs) to generate descriptive prompts. This unique integration enables the creation of 3D visuals that emphasize tourism-related aspects, offering a new dimension of engagement for users exploring virtual tourist destinations.

- **Effective Use of NLP:** By leveraging NLP to generate descriptive prompts, we successfully bridge the gap between image content and 3D visualization. This step enhances the relevance and accuracy of the 3D models, aligning them with specific tourism interests and providing a more targeted user experience.
- **High-Quality 3D Visuals:** Our system effectively generates high-quality 3D visuals that highlight key features of tourist locations. The use of advanced 3D rendering techniques ensures that the visuals are both detailed and immersive, making them suitable for applications in virtual tourism.
- **Real-World Applicability:** The project demonstrates significant potential for real-world applications in the tourism industry. By offering an interactive and visually appealing way to explore tourist destinations, our system can enhance travel planning and virtual exploration experiences.

2.9 Future Work

While the "Virtual 3D Tour" project has achieved its primary objectives, several areas for future work and improvement have been identified:

- **Expansion of Dataset:** To improve the accuracy and diversity of 3D visualizations, expanding the dataset to include a wider variety of tourist destinations and image types is essential. Incorporating more comprehensive datasets will enhance the system's ability to generate detailed and realistic 3D models.
- **Enhanced NLP Integration:** Future work could involve refining the NLP component to handle more complex and nuanced descriptions. Improving the model's ability to generate contextually relevant prompts will lead to even more accurate and engaging 3D visuals.
- **Real-Time Processing:** Implementing real-time processing capabilities would allow users to generate 3D visuals on-the-fly from live image inputs. This enhancement could significantly improve the user experience and broaden the potential applications of the system.
- **User Feedback and Interaction:** Incorporating user feedback mechanisms and interactive features could further enhance the system. Allowing users to customize and interact with the 3D visuals in real-time would make the application more engaging and user-friendly.
- **Integration with Augmented Reality (AR):** Exploring integration with AR technologies could provide a more immersive experience by overlaying 3D visuals onto

real-world environments. This integration would bridge the gap between virtual and physical experiences, offering innovative applications in tourism and beyond.

Bibliography

[1] A Kolyshkin and S Nazarovs. Stability of slowly diverging flows in shallow water. *Mathematical Modeling and Analysis*, 2007.