Project Design Phase-I

Solution Architecture

Date	23 October 2023
Project Name	Project - Image Caption Detection
Maximum Marks	4 Marks

Solution Architecture:

VisualClarity's solution architecture is designed with a modular and scalable approach. At its core lies an image interpretation module, employing cutting-edge Convolutional Neural Networks (CNNs) for accurate object and scene recognition within uploaded images. The interpreted data then flows into a Natural Language Generation module, where advanced Natural Language Processing (NLP) models, including Recurrent Neural Networks (RNNs) or Transformers, generate descriptive and contextually relevant textual captions. These components work seamlessly together, underpinned by scalable cloud infrastructure, ensuring real-time responsiveness and adaptability to increasing demand. The user-facing layer comprises a user-friendly interface designed for accessibility, compatible with screen readers and voice commands. This layer connects users to the backend modules, providing visually impaired individuals with immediate and meaningful image descriptions. Continuous improvements and user feedback loops are integrated, refining the architecture and enhancing the overall user experience.

Solution Architecture Diagram: