

Amrita School of Computing, Amritapuri

Department of Computer Science and Engineering (AI)

B. Tech 2021-2025 CSE(AI)

**“Deep Vision: Automated Image Captioning and
Multi-Modal Knowledge Graph Construction”**

Group No: AI_AB2

Team Members

Sayanthika K R - AM.EN.U4AIE21058

Adithya R K Nambiar - AM.EN.U4AIE21105

Alap S Suresh - AM.EN.U4AIE21110

Gokul Krishna B R - AM.EN.U4AIE21131



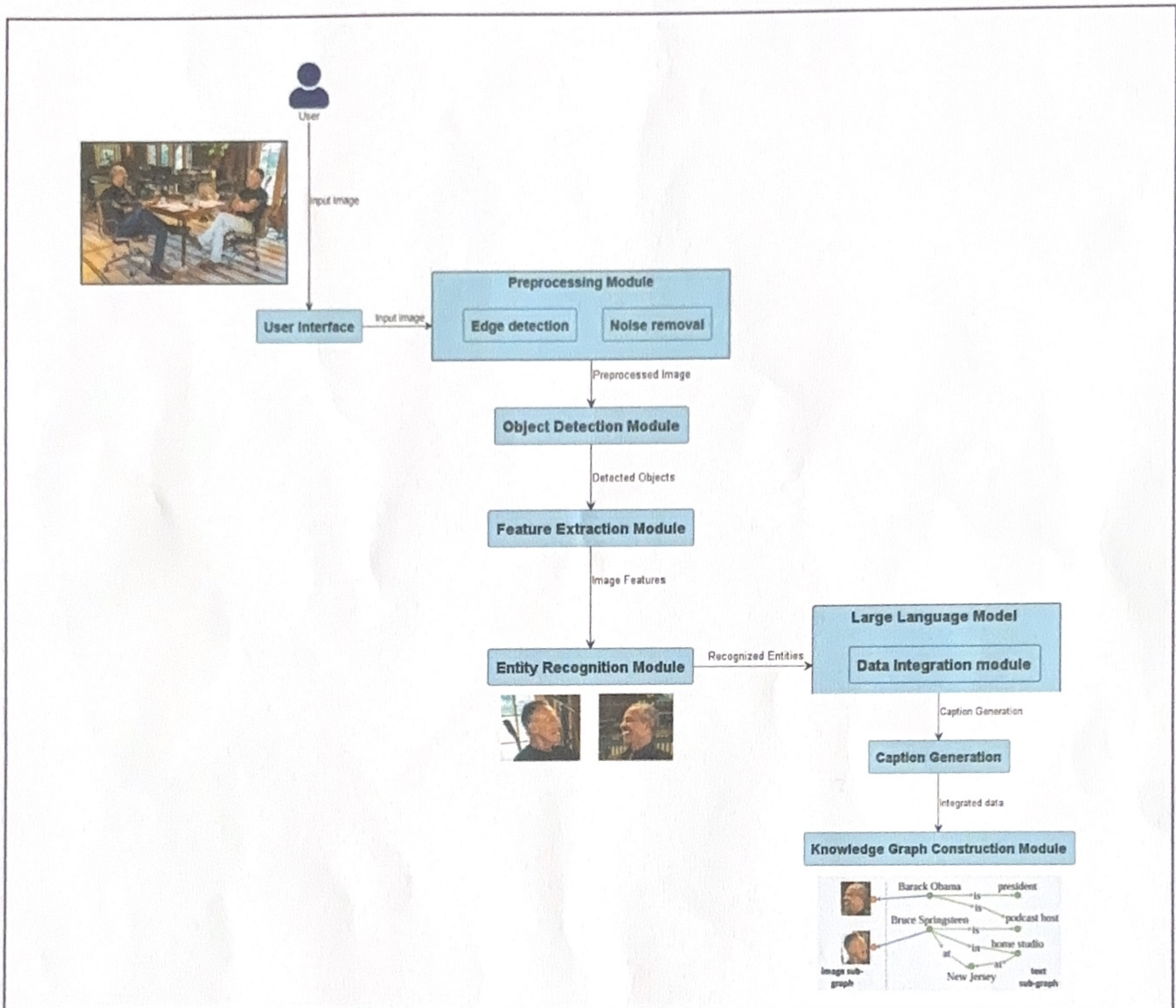
August 2024

GUIDE
Dr. Lekshmi S Nair

COORDINATOR
Ms. Raji Ramachandran

COORDINATOR
Ms. Nisha K S

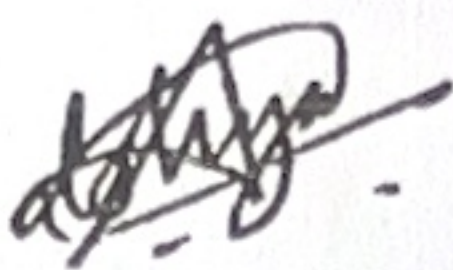

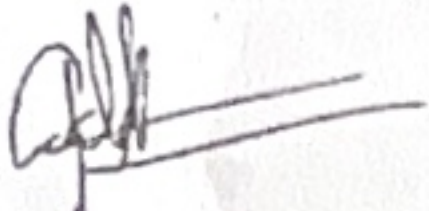
Block Diagram



Summary

The system follows a multi-stage process, starting with the preprocessing of the input image using techniques like edge detection and noise removal to enhance image quality. After preprocessing, the image is passed through an object detection module to identify key objects. Next, the feature extraction module analyses the detected objects to capture their significant attributes, such as shape and colour. The entity recognition module identifies and classifies these objects into meaningful entities. A large language model (LLM) then generates descriptive captions based on the recognized entities, while the data integration module organizes these entities into a knowledge graph, linking them with their relationships for contextual understanding.

Students' Name and Signature

Student's Name	Signature
Sayanthika K R	
Adithya R K Nambiar	
Alap S Suresh	
Gokul Krishna B R	

Guide's Name and Signature

Guide's Name	Signature
Dr. Lekshmi S Nair	

Date: 05/09/2024