Project Documentation – Jio Hotstar AdVison & Analytics

Problem Statement:

**Project Problem Statement: Brand Detection and RAG Chat Dashboard**

**1. Introduction**

In the modern digital era, video content is ubiquitous across various platforms such as social media, advertising, and surveillance. Companies and content creators often need to analyze videos to detect and track the presence of brand logos for purposes like brand monitoring, advertising effectiveness, and competitive analysis. However, manually scanning through hours of video footage to identify brand appearances is time-consuming and prone to human error.

**2. Problem Definition**

There is a pressing need for an automated system that can:

* Detect and identify brand logos in video content with high accuracy.
* Store the detection data in a structured format for further analysis.
* Provide an intuitive interface for users to upload videos, view detection results, and interact with the data through natural language queries.

**3. Project Objectives**

The main objectives of this project are:

* **Automated Brand Detection**: Utilize a pre-trained or custom YOLO (You Only Look Once) model to detect and identify brand logos in uploaded videos.
* **Data Storage**: Store detection metadata (such as video name, frame number, timestamp, detected logo name, confidence score, and bounding box coordinates) in a PostgreSQL database.
* **Dashboard Interface**: Develop a Streamlit-based dashboard that allows users to:
  + Upload video files for processing.
  + View real-time video processing with bounding boxes around detected logos.
  + Access a database explorer to review stored detections.
  + Visualize analytics and statistics about the detections (e.g., brand frequency, confidence distribution).
  + Interact with a RAG (Retrieval-Augmented Generation) AI chatbot that answers natural language questions about the detection data.

**4. Scope**

The project will focus on:

* Supporting common video formats (MP4, AVI, MOV, MKV).
* Using a YOLO model (initially a pre-trained model for common objects, with the option to use a custom-trained model for specific brands).
* Storing data in a PostgreSQL database and providing functionality to initialize the database schema.
* Providing a web-based dashboard with multiple tabs for video processing, AI chat, analytics, and data exploration.
* Integrating with Groq AI for the RAG chatbot to answer questions based on the database content.
* Optional: Uploading videos and results to AWS S3 for cloud storage.

**5. Functional Requirements**

1. **Video Upload**: Users can upload video files through the dashboard.
2. **Video Processing**: The system processes the video frame by frame (or at a specified interval) to detect brand logos.
3. **Real-time Display**: During processing, the video is displayed with bounding boxes and labels for detected logos.
4. **Data Storage**: Detection results are stored in the PostgreSQL database.
5. **Database Management**: The system includes functions to initialize the database and clear existing data.
6. **Analytics Dashboard**: Visualizations of brand detection statistics, including frequency and confidence levels.
7. **AI Chat Interface**: A chatbot that uses the Groq API to answer questions about the detected brands.
8. **Data Export**: Users can download detection results as CSV files.

**6. Non-Functional Requirements**

1. **Performance**: The system should process videos efficiently, leveraging the YOLO model for fast object detection.
2. **Usability**: The dashboard should be user-friendly and intuitive.
3. **Scalability**: The system should be able to handle multiple video uploads and process them in a reasonable time.
4. **Reliability**: The database operations and video processing should be robust and handle errors gracefully.

**7. Technical Stack**

* **Backend**: Python, Streamlit
* **Computer Vision**: YOLO (Ultralytics)
* **Database**: PostgreSQL
* **AI Chat**: Groq API
* **Cloud Storage**: AWS S3 (optional)
* **Data Visualization**: Plotly

**8. Challenges**

* **Model Accuracy**: Ensuring the YOLO model is accurately trained to detect a wide range of brand logos.
* **Processing Speed**: Balancing between processing speed and detection accuracy, especially for long videos.
* **Data Management**: Efficiently storing and retrieving large amounts of detection data.
* **Natural Language Understanding**: The RAG system must correctly interpret user questions and provide accurate answers based on the database.

**9. Expected Outcomes**

* A fully functional web application that automates brand logo detection in videos.
* A structured database of detection results.
* An interactive dashboard for visualization and querying of the data.
* A RAG-based chatbot that provides insights about the detected brands.

**10. Conclusion**

The Brand Detection and RAG Chat Dashboard project aims to streamline the process of brand monitoring in videos, providing a powerful tool for marketers, content creators, and analysts. By leveraging cutting-edge technologies in computer vision and natural language processing, the project will deliver an efficient and user-friendly solution.