India's No.1 E-School Challenge is back!



SHAPING INDIA'S TECHSCAPE,

About Team

Title: Smart Vision System + Team Sentinels (Robotics Challenge)

Team Name: Sentinels

Team Members:

- Hrithik Reddy
- Rishabh Sharma
- Aditya Dargan
- Ayush Kumar Singh Rathor
- T M Adi Venu Gopala Reddy Padala

College/University:

- Netaji Subhas University of Technology, Delhi: Rishabh Sharma, Ayush Kumar Singh Rathor
- **Delhi Technological University, Delhi:** Aditya Dargan
- SRM Institute of Science and Technology: Hrithik Reddy, T M Adi Venu Gopala Reddy Padala
- Date: 20th October 2024

Executive Summary

The Smart Vision System leverages AWS cloud infrastructure and machine learning models to provide real-time video processing, image classification, object detection, object tracking, OCR and Entity Extraction. The system is designed to securely stream video from a conveyor belt, analyze the footage for key insights, and display the results to operators through an intuitive dashboard

1. **Input:** Video streams from an on-premises camera monitoring a conveyor belt are securely transmitted to the cloud via a VPN gateway.

2. Processing:

- Backend services running on EC2 instances (FastAPI) handle video processing.
- Machine learning algorithms such as YOLOv11, MobileNetV2, Llama3.2-11B-vision and DeepSort are used for object detection, classification, OCR, object tracking and entity extraction.
- 3. **Storage:** Processed images are stored in S3 buckets, while structured data such as object tracking logs and entities are stored in DynamoDB.
- 4. **Monitoring:** AWS CloudWatch is used to log system metrics and ensure smooth operation.
- 5. **Output:** Operators interact with the processed data and video streams through a ReactJS frontend hosted on AWS Amplify.

Technical Approach

Teams to submit a solution proposal ppt that includes details on the technology used, hardware specifications, items used to train the model, and code, accompanied by a video simulation

1. Technology Stack:

- a. Cloud Platform: AWS VPC, EC2 instances, S3 buckets, DynamoDB, AWS Bedrock ,AWS Sagemaker, AWS Amplify
- b. **Backend Framework:** FastAPI for processing video and managing backend services(Server Sent Events for real time processing)
- c. **Frontend Framework:** ReactJS for the operator dashboard, hosted using AWS Amplify.

2. Hardware Specifications:

- a. **Camera:** High-definition camera mounted on the conveyor belt to capture real-time footage.
- b. **Compute Power:** EC2 instance and AWS Sagemaker GPU support (e.g., p3.2xlarge) for real-time video processing.
- c. Storage:
 - i. S3 Buckets for image storage and long-term archival.
 - ii. DynamoDB for storing logs and processed data entries.

3. Machine Learning Models:

- a. YOLOv11: Used for object detection to identify items on the conveyor belt.
- b. **MobileNetV2:** Used for classification tasks, such as determining object freshness or quality.
- c. **DeepSort:** Used for object tracking to precisely track the objects on conveyor belt without duplicate entries
- d. Llama3.2-11b-vision: Used for OCR and entity extraction from the detected objects

Technical Approach:

Teams to submit a solution proposal ppt that includes details on the technology used, hardware specifications, items used to train the model, and code, accompanied by a video simulation

4) Training Data:

- a. **Dataset:** Imitated Conveyor belt and collected data manually for object detection, used the kaggle Fresh Fruits and Vegetable dataset for classification
 - i. https://app.roboflow.com/flipkart-grid-gqtdk/belt-nublw/2
 - ii. https://www.kaggle.com/datasets/swoyam2609/fresh-and-stale-classification
- b. **Annotations:** Manually annotated data for object detection in Roboflow annotator.
- c. Training Tools:
 - i. Kaggle P100 GPU was used for training/fine-tuning the models used in the system.
 - ii. Transfer Learning: Pre-trained models (e.g., MobileNetV2 and YOLOv11) fine-tuned with the custom dataset.

5) Simulation Video:

A simulation of the system demonstrates how the camera captures real-time footage, how the backend processes the video, and how the operator dashboard displays the results

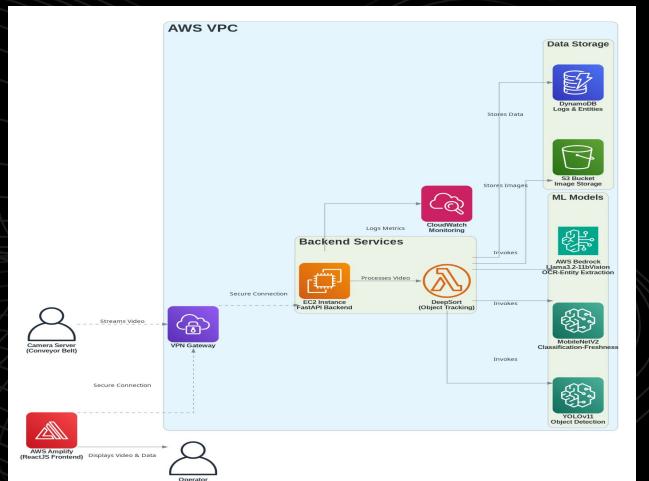
6) Source Code:

https://github.com/Hrithik2212/Team-Sentinel

<u>Video Link</u>: https://www.youtube.com/watch?v=Z_sjBRmIVX8

ARCHITECTURE

NOTE: Set Quality to 1080 for the video in the next slide, If the video is not available click the link in the previous tab



Smart Vision System Architecture



India's No.1 E-School Challenge is back!

Flipkart 6.0

SHAPING INDIA'S TECHSCAPE



Thank You