

Real-Time GIKI Campus Building Detection and Instance Segmentation using YOLOv8

Problem Statement

The goal of this project is to develop a real-time AI system capable of detecting and segmenting campus buildings from video footage. This includes identifying each building by name, differentiating between multiple structures, and generating both bounding boxes and segmentation masks for visualization.

Dataset Description

Source: Custom dataset collected by recording videos on GIKI campus from different angles, buildings, and lighting conditions.

Size: 116 images

Train: 81 images (70%)

Validation: 23 images (20%)

Test: 12 images (10%)

Classes (14 total):

HLB_bank, Main_gate, auditorium, central_library, clock_tower, faculty_of_Basics_Science, faculty_of_electrical_and_computer_science, faculty_of_material_and_chemical_eng, faculty_of_mechanical_eng, hostel_11, hostel_12, laundry_shop, new_academic_block, tuck

Annotation:

Original annotations done in Roboflow for object detection (bounding boxes). For segmentation, bounding boxes were converted into rectangular polygon masks to allow YOLOv8-seg training without re-labeling.

Methodology

Data Collection

Recorded videos across campus capturing all target buildings. Extracted frames and uploaded them to Roboflow.

Annotation

Bounding boxes were drawn for each building using Roboflow's labeling tool. Dataset exported in YOLOv8 format. Bounding boxes converted to rectangular masks for segmentation.

Model Training

Object Detection (YOLOv8 Detection)

Model: yolov8s.pt pretrained on COCO dataset.

Training configuration:

Epochs: 50

Image size: 640×640
Batch size: 8
Optimizer: Adam (default)
Output: Bounding boxes around detected buildings with class labels.
Instance Segmentation (YOLOv8-Segmentation)
Model: yolov8s-seg.pt pretrained on COCO.
Training configuration:
Same as detection, but with converted polygon masks.
Output: Bounding boxes + mask overlays.

Results

Detection Model

Successfully detected buildings in both images and video.
Saved output videos with bounding boxes for all 14 classes.

Segmentation Model

Generated rectangular masks around buildings along with labels.
Worked well for visual separation of multiple buildings in the same frame.

Evaluation

Due to small dataset size, metrics like mAP were not primary focus — qualitative visual performance was emphasized.

Visual inspection confirmed:

Correct class labeling in most frames.

Minimal false positives.

Segmentation masks accurately matching bounding box shapes.

Insights

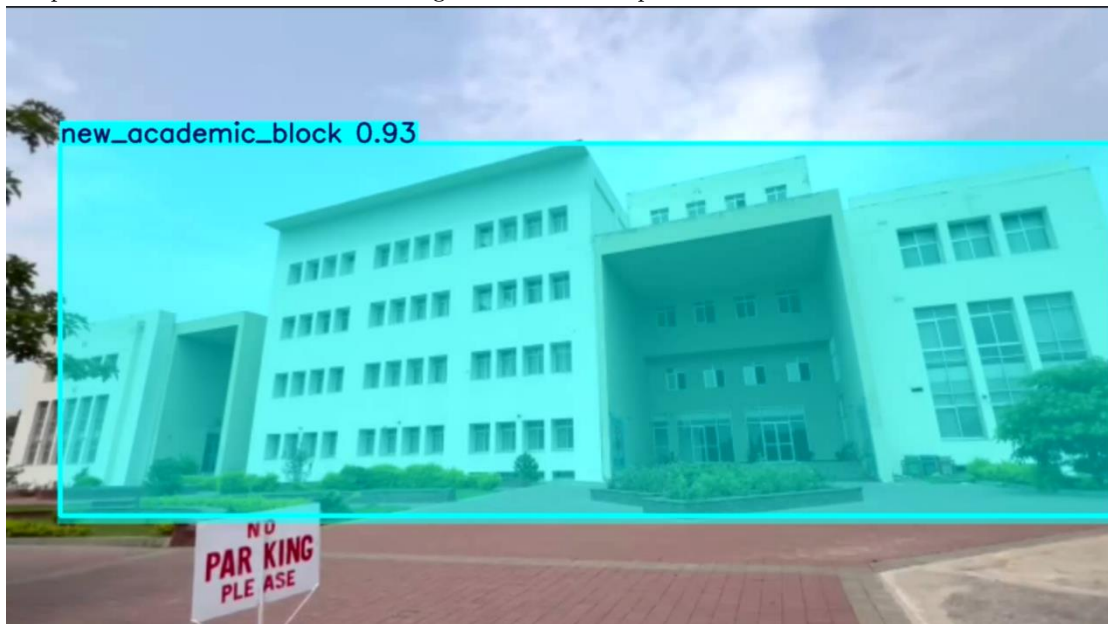
YOLOv8 detection worked very well even with limited data, likely due to strong transfer learning from COCO weights.

The segmentation masks (rectangle-based) are not true shape masks, but still improve visual clarity in multi-object scenes.

Dataset expansion with real polygon masks could significantly improve mask quality.



Comparison of detection vs segmentation outputs



Conclusion

We successfully created and trained two AI models:

YOLOv8 Object Detection — for bounding box detection of 14 campus buildings.

YOLOv8 Instance Segmentation — for mask overlay visualization.

This system can be integrated into a real-time campus guide app, allowing visitors or students to identify buildings instantly via camera input.