HALS

Bhavana Rao

Sriram Kothapalli

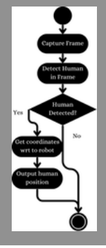
Objective

Designing a system to detect and track one or more humans using a monocular video camera.
Output location information should be directly usable in a robot's reference frame.
Key features: Real-time detection, localization in robot's coordinate system.

Technical Approach

YOLOv5, a classical computer vision model is used for detecting humans. The YOLO network consists of a CNN backbone, a neck that combines image features and a head that puts bounding box over the detected object.
Geometric Computer Vision is used for localization of the human position wrt to robot's coordinates.

Class Diagram



Timeline

- Phase 0: Proposal
- Phase 1: Plan - TDD to guide implementation and unit tests. Implement GitHub CI and CodeCov.
- Phase 2: Implementation - Final running code, Doxygen documentation, Product Backlog