

Human



Faster-
Whisper
(STT)

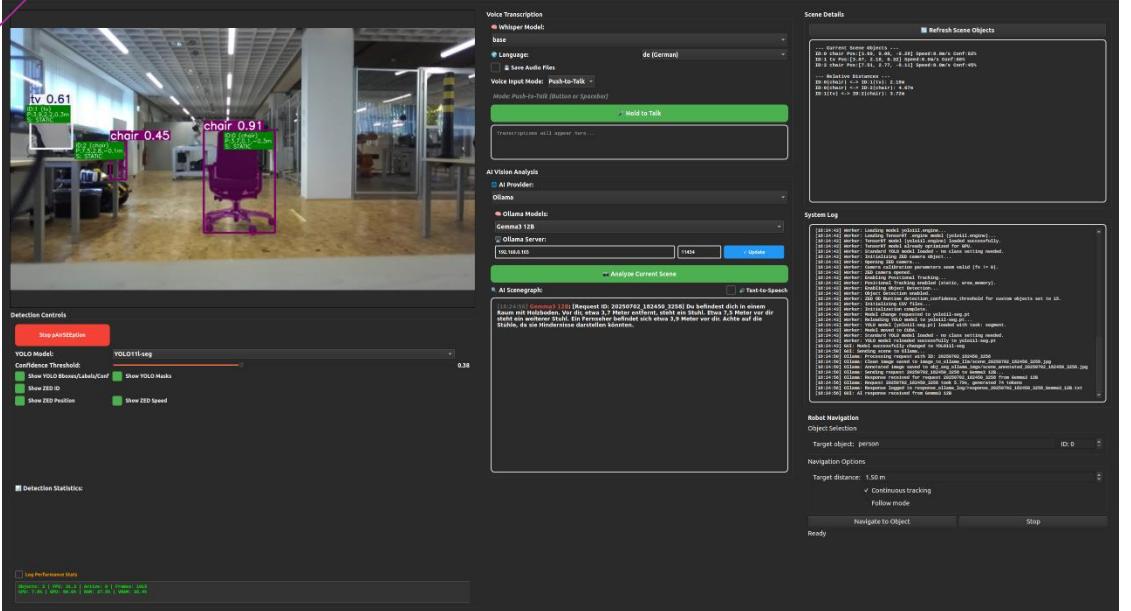
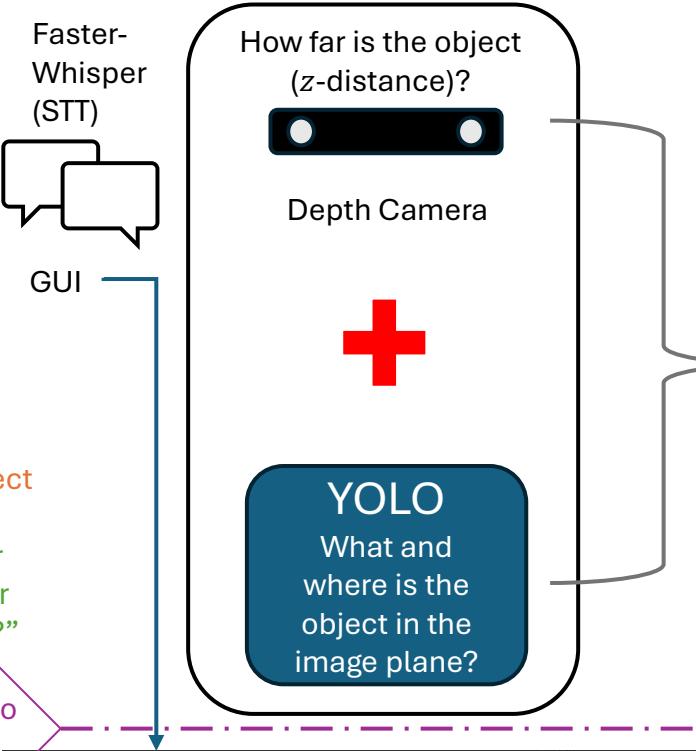


GUI

Possible user input:

- a. "Analyze the scene!"
- b. "Drive to object ABC!"
- c. "What is your current motor temperature?"
- d. "Can you explain how to use your system?"

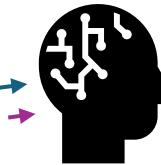
Object detection and localization



Intent analyzer

Hybrid voice assistant (TTS: VITS)

LLM/VLM answer



a. "There is a car in front of you, 5 meters away."

d. "Sure, you can give me voice commands to analyze the scene for a blind person, for example, or you can tell me a navigation destination in my field of vision that I should drive to, or I can follow any recognized object by YOLO!"

b. "I will drive to object ABC.... Destination reached!"

c. "My left motor temperature is 25 °C and my right motor temperature is 26 °C."

Write and subscribe to ROS2 topics to drive or query status

ROS2

