

GONK (Generative OpenSource Neural Kompanion).

GONK (Generative OpenSource Neural Kompanion)

A project in IN5590 by Alan Handren-Hashim Aziz and Kasper André Meldre Bilet
E-mail: ahaziz@ifi.uio.no and bilet@ifi.uio.no

1 Intro

GONK is the first ever Generative OpenSource Neural Kompanion straight from Goop. It has the ability to detect and converse with humans, display different moods and move its upper body.

Goals:

- Arm and torso movement
- Human recognition
- Conversation ability

Nice to have's:

- Mimicking human movements (Not achieved)
- Display emotions (Achieved)

2 Tools

Tools used for this project were as follows:

- Dynamixel AX12
- Dynamixel AX18
- Raspberry Pi 3 B+
- Raspberry Camera Module 3
- 4 Ohms Speaker
- MAX98357A
- Teensy 4.0
- Sparkfun: Electret Microphone
- Sparkfun: Color 320x240 TFT Touchscreen, ILI9341 Controller Chip

YOLOv11 was used for human detection. To make it run more smoothly on the Raspberry Pi 3B+, the model was exported to the NCNN format made for mobile devices.

A LLM pipeline using OpenAI, ElevenLabs and Microsoft Azure backend was also used to lessen the computational load on the Raspberry Pi 3B+ and make the conversation more organic.

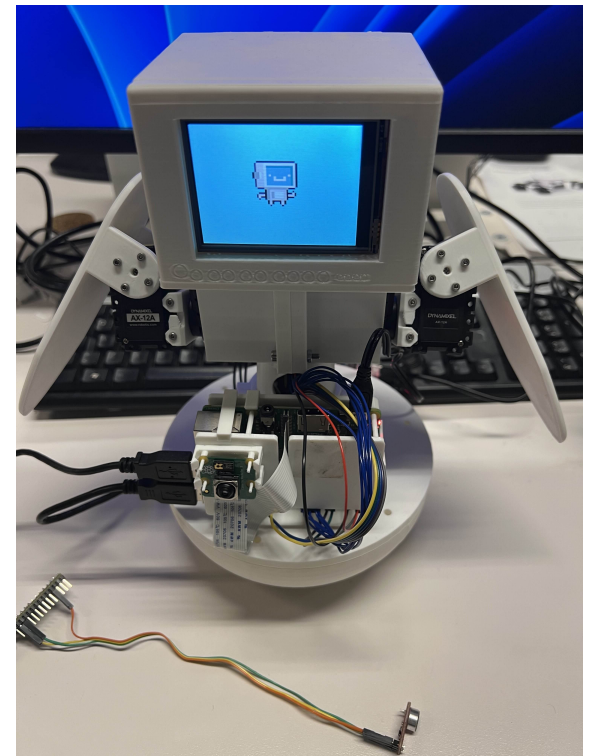
3 Conclusion

Our goals were all achieved, and we also managed to achieve one of our nice-to-haves. However, we were not able to implement human movement mimicking.

Using an NCNN instead of the standard YOLOv11.pt weights or ONNX made it less computationally heavy for the Raspberry Pi 3B+, making it run faster than ONNX when both were tested. Setting up an LLM pipeline using Microsoft Azure, OpenAI, and ElevenLabs also lessened the computational load on the Raspberry Pi 3B+ by running it on their servers instead of natively.

A big part of what could have been done differently is to make the body of the robot larger to improve cable management. The current design has enough space for the current cables, but it is a tedious task to put them together. The robot was supposed to be modular and have easy-to-replace parts if needed, but in its current form, this would prove challenging. The design would need to be bigger and changed to a more modular design.

Extra figures



UNIVERSITY
OF OSLO

GitHub repo →

