CREATIONS de UiTM

INTERNATIONAL MEGA INNOVATION CARNIVAL 2024

Navigating Innovation and Seizing Global Fortune

CHANGE THE WORLD THROUGH INNOVATION

e-PROCEEDING

27th APRIL 2024

UNIVERSITI TEKNOLOGI MARA
CAWANGAN SELANGOR, KAMPUS DENGKIL
MALAYSIA

ORGANISED BY:

Universiti Teknologi MARA
Pusat Asasi

EYE (i) ROBOT: AI-Driven Approach to Home Assistants

*Muhammad Safrul Rosli and Raudzatul Fathiyah Mohd Said

Centre of Foundation Studies, Universiti Teknologi MARA, Cawangan Selangor, Kampus Dengkil, 43800 Dengkil, Selangor, Malaysia

*Corresponding author: muhdhdsafrul@gmail.com

ABSTRACT

A technological invention, called smart home assistant, originally held the potential to improve quality of life. However, the idea was long buried by the masses due to its underlying unreliability. Hence, EYE (i) ROBOT serves as an AI-Driven solution to home assistants. Feature includes automatic speech recognition, face recognition, home automation and home monitoring. Made to offer a better solution than existing home assistants in the market. The goal is to create a home assistant that is reliable, natural-like, and capable of securing private spaces. This can be achieved by utilising various Artificial Intelligence (AI) technologies. Findings shows that EYE (i) ROBOT is able to identify humans. Moreover, it is able to control a smart light within reasonable timeframe. That being said, EYE (i) ROBOT is targeted for building owners whether it's a house or office, seeking security and assistance within their private spaces. In conclusion, EYE (i) ROBOT is a further step towards our mission to benefit humanity.

Keywords: artificial intelligence; raspberry pi; python.

1. INTRODUCTION

The scope of this paper is to improve the existing smart home assistants by implementing latest AI technologies. The objective is to facilitate a more natural and intuitive interaction between users and EYE (i) ROBOT. Recent review of pertinent literature has shed light in several key areas of focus in enhancing user interaction and experience with smart home assistants. These include natural language processing (NLP), computer vision, context-aware interaction paradigms, and human-robot interaction (HRI) research. The advancements mentioned greatly contributes to our mission. To ensure fluidity in development, the methodology involves Requirements Analysis, Software Architecture Design, Development and Testing, Hardware Design and Integration, and lastly Testing and Evaluation. Initial evaluation of EYE (i) ROBOT demonstrates promising results. It is able to respond within a reasonable timeframe and exhibits high accuracy in recognizing faces.

2. METHODOLOGY

This methodology outlines the systematic approach taken in EYE (i) ROBOT development process. The system comprises both front-end and back-end features written in Python (Python 3.11). It is housed within a meticulously designed casing with carefully selected internal components.

2.1. Requirement analysis

Detailed requirements for software and hardware components were gathered to establish clear objectives and constraints. Key functionalities including facial recognition, speech recognition, home automation, and language model processing were identified for the back-end, while front-end features such as simulated 'awareness' through a digital eyeball, user-to-robot text display and robot-to-user text display were delineated.

2.2. Software architecture design

The software architecture was designed for a seamless join between back-end and front-end features. Furthermore, modular design principles were applied to ensure scalability, maintainability, and efficiency. Each module was defined to encapsulate specific functionalities, promoting ease of development and future enhancements.

2.3. Development and testing

Back-end features were developed iteratively, utilising advanced algorithms and models for tasks such as facial recognition, speech recognition and Generative AI integration. Rigorous testing were conducted to validate functionality, reliability, and performance. Front-end features were implemented concurrently, with careful attention to user experience.

2.4. Hardware design and integration

The casing is 3D printed out of PETG+ filament and designed in FreeCAD. The internal components such as Raspberry Pi 4B, IPS display, and audio I/O system, were selected to complement the software functionalities.

2.5. Testing and evaluation

Hardware components were assembled and the software is deployed on-device. Several testing was then performed to solicit benchmarks to ensure compatibility, functionality, and reliability before go-live.

3. RESULTS AND DISCUSSION

3.1 Speech comprehension proficiency

EYE (i) ROBOT exhibits a more nuanced understanding compared to traditional smart home assistants. Through advanced speech recognition algorithms and natural language processing capabilities, the robot accurately interprets user commands, enabling seamless interaction. As of this writing, Gemini is chosen as the LLM-model. In tabulated benchmark (Table 1), Gemini's Pro model achieved comparable but slightly inferior to OpenAI's models. However, Gemini is more cost effective.

Table 1. Model Benchmark (Akter, S. N., et al., 2023)

	Model				
Task	**Dataset**	**Gemini Pro**	**GPT 3.5 Turbo**	**GPT 4 Turbo**	**Mixtral**
Knowledge-based QA	MMLU (5 shot)	65.22	67.22	**80.48**	68.81
	MMLU (CoT)	62.09	70.07	**78.95**	59.57
Reasoning	BIG-Bench-Hard	67.53	71.02	**83.90**	60.76
Mathematics	GSM8K	76.42	78.01	**92.72**	71.65
	SVAMP	81.10	82.30	**92.60**	81.60
	ASDIV	85.31	89.07	**92.75**	83.16
	MAWPS	96.50	98.00	**98.67**	96.00
Code Generation	HumanEVAL	59.76	74.39	**76.83**	45.12
	ODEX	39.86	**52.62**	45.79	40.55
Machine Translation	FLORES	53.31	52.43	**54.00**	40.97
	Unblocked	21.68	40.00	**48.24**	30.27
	FLORES All				
Web Agents	WebArena	7.12	8.87	**14.90**	1.39

3.2 Responsiveness

EYE (i) ROBOT demonstrates the ability to recognize users with remarkable precision and speed. It consistently identifies faces even under varying lighting conditions and facial expressions.

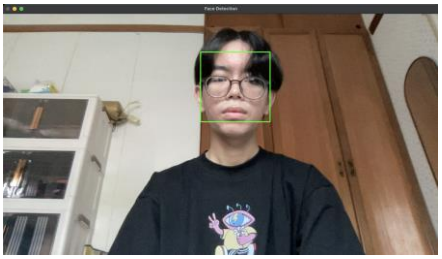

Figure 1. Face Detection

3.3 Robustness of software

EYE (i) ROBOT demonstrates the ability to recognize users with remarkable precision and speed. It consistently identifies faces even under varying lighting conditions and facial expressions.

Table 2. Tkinter and PyQt Performance Benchmark

Framework	CPU Usage (%)	RAM Usage (MB)
Tkinter	4.6%	1540.0
PyQt	2.8%	184.0

4. CONCLUSION

EYE (i) ROBOT is a further step towards our mission to benefit humanity. However, despite the capabilities shown, we should note that there are limitation to the use of these AI technologies. There is a continued need for ongoing research and development in "hallucinations" generated. Be that as it may, we are enthusiastic to push towards our broader future to develop the best AI-assistant.

REFERENCES

Akter, S. N., Yu, Z., Muhamed, A., Ou, T., Bäuerle, A., Cabrera, N. A., Dholakia, K., Xiong, C., & Neubig, G. (2023, December 18). *An In-depth Look at Gemini's Language Abilities*. arXiv.org. https://arxiv.org/abs/2312.11444