RESEARCH ARTICLE OPEN ACCESS

DEVELOPMENT OF A SMART FACE BOT WITH VOICE OUTPUT AND DYNAMIC FACIAL EXPRESSIONS

S. Annamalai[1], R.V. Rahul[2], R. Manoj Kumar[3], A. Sachin Deepak[4]
Mrs.M. Devi [5], Lecturer
Electronics (Robotics), Murugappa Polytechnic College, Sathyamurthy Nagar, Avadi, Chennai 600062.
[1]Email: annamalaisanguabani@gmail.com
[2]Email: rrv92550@gmail.com
[3]Email: manojkumar845831@gmail.com
[4]Email: deepaksachin64@gmail.com

--************************----------------------------------

Abstract:

A Smart Face Bot is a robotic system powered by an ESP32 microcontroller, controlling facial expressions and speech. It uses a PCA9685 module to drive servo motors for mouth and head movements. An OLED display represents dynamic facial features like eyes. A PAM amplifier and speaker enable speech synthesis for interactive communication. The bot generates vocal output using text-to-speech (TTS) technology. Servo-controlled mouth movements sync with speech for realism. The OLED display enhances expressiveness by changing eye shapes. This system creates an engaging, expressive robotic face with lifelike interactions.

***Keywords**:* ESP32; Microcontroller; OLED; PCA9685; TTS; Bot; pam Amplifier

--************************----------------------------------

I.INTRODUCTION

The Smart Face Bot is an interactive robotic system designed to express emotions and communicate through speech. It is powered by an ESP32 microcontroller, which controls various components to create lifelike interactions. The bot features an OLED display that dynamically represents eyes, enhancing its expressiveness.

A PCA9685 module is used to control multiple servo motors, enabling smooth and synchronized mouth and head movements. The mouth movements are precisely coordinated with speech to create a realistic talking effect. A PAM amplifier powers the speaker, ensuring clear and audible voice output.

Using text-to-speech (TTS) technology, the bot can generate spoken responses, making it more interactive. The OLED display allows the bot to change eye expressions, adding personality to its communication. The ESP32 processes inputs and controls responses efficiently, making it a compact and versatile system.

This robotic face can be used in AI assistants, educational tools, or entertainment applications. It enhances user engagement by responding with both speech and facial animations. The combination of visual and auditory interaction makes it more engaging than a simple speaker-based assistant.

II. METHODOLOGY

ESP32 Dev Kit:

The ESP32 microcontroller serves as the brain of the Smart Face Bot, controlling all components, including the OLED display, servo motors, and audio output. It processes commands, manages communication between peripherals, and generates speech output using text-to-speech (TTS) technology.

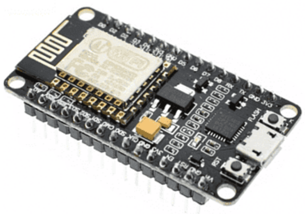

Fig 1. Esp32 Dev Module

OLED Display (128x64, I2C):

The OLED display is used to show facial expressions, such as blinking, happiness, or sadness. The ESP32 sends preloaded PNG/JPEG images to the display based on the robot's speech or user interaction.

Fig 2. OLED Display

By integrating robotics and AI, the Smart Face Bot provides an advanced way to interact with machines. Its smooth movements, expressive display, and clear speech make it a compelling platform for interactive communication.

PCA9685 Servo Driver:

Since ESP32 has limited PWM pins, the PCA9685 module is used to control multiple servo motors efficiently. It communicates with ESP32 via I2C, allowing precise angle adjustments for mouth movement, head rotation, and other robotic motions.

Fig 3. PCA 9685 Module

Servo Motors:

The MG995N servo motor is a high-torque digital servo used in the Smart Face Bot for jaw movement, head rotation, and structural support. It operates at 4.8V to 7.2V, providing a torque of up to 10 kg/cm, making it ideal for handling heavier 3D-printed parts. With a metal gear system, it ensures durability, precision, and smooth motion control.

Fig 4. Servo Motor

PAM8403 Audio Amplifier & Speaker:

The ESP32's built-in DAC (Digital-to-Analog Converter) is not strong enough to produce loud audio output. The PAM8403 amplifier boosts the sound signal, and the speaker outputs clear and audible speech.

Fig 5. PAM8403 Audio Amplifier & Speaker

Power Supply:

ESP32 operates at 5V via USB or an external power Module.Servo motors require a separate 6V–7.4V power source to prevent overloading ESP32.

Fig 6. Power Supply

III. SOFTWARE IMPLEMENTATION:

OLED Display for Facial Expressions:

The OLED display is used to show various facial expressions, including neutral, happy, sad, and angry, by displaying preloaded PNG/JPEG images**.** These expressions change dynamically based on speech output and user interaction. The Adafruit SSD1306 and GFX libraries handle the display operations, allowing efficient rendering of facial images on the 128x64 OLED screen.

Servo Motor Control (Mouth and Head Movement):

The PCA9685 module is used to control multiple servo motors**,** ensuring precise mouth and head movements. The mouth movement is synchronized with speech output, making the bot appear more expressive when speaking. Additionally, predefined head positions allow the bot to perform nodding, turning, and tilting motions, enhancing its interactive capabilities

Speech Synthesis Using TTS:

The text-to-speech (TTS) engine converts text input into speech, allowing the bot to communicate with users. Depending on the implementation, either Google-based TTS or ESP32's built-in TTS engine is used to generate voice output. The PAM8403 amplifier boosts the audio signal from the ESP32, ensuring clear and loud speech through the 8Ω speaker.

Integration of Facial Expressions and Speech:

The software integrates facial expressions with speech output, making the bot appear more lifelike. When the bot speaks, it selects appropriate facial expressions based on the context. For example, when saying "Hello!"**,** the bot displays a happy face, and when saying "I don't know", it shows a confused expression**.** This dynamic synchronization improves the bot's ability to engage with users effectively.

Wireless Control and Future Enhancements

The ESP32's built-in Wi-Fi and Bluetooth capabilities allow remote control and real-time updates. Users can send text commands wirelessly via a mobile app or web interface**,** which the bot converts into speech and expressions. Future enhancements may include AI-based emotion detection, allowing the bot to adjust its expressions based on user emotions, making interactions more natural and engaging.

IV. STEP-BY-STEP DEVELOPMENT OF A SMART FACE BOT

Step 1: 3D Printer Parts

We have used 3D-printed parts for each component of the robot. I chose a red and white color scheme for the robot. We printed all the parts using an Ender 3 V2 3D printer with basic Cura settings, including 20% infill and a 0.2 mm nozzle. For a better finish, the 3D-printed parts can be smoothed using sandpaper.

Use M3 screws to attach the servo horns to the legs, and assemble the hands in the same manner. Ensure the servo angles are correctly adjusted to make the robot stand straight and steady.

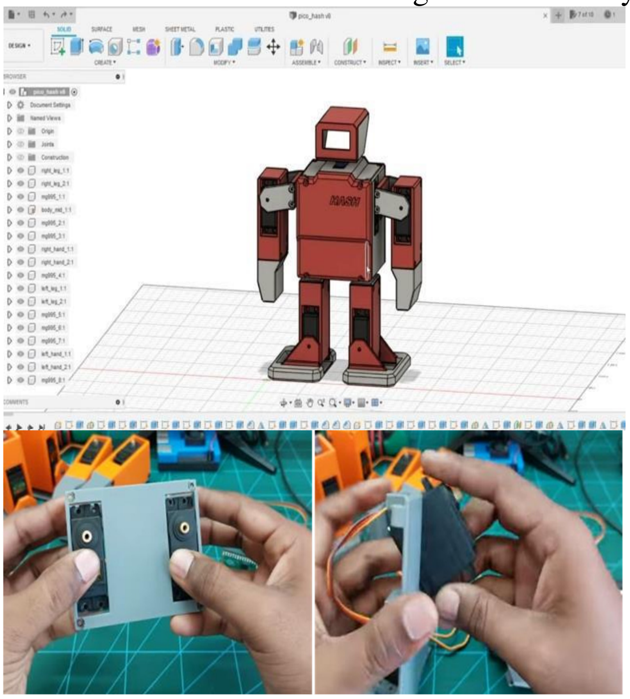

Fig 8. Hand & Leg Assembly

Step 3: Assembling the Body

Similarly, assemble the servos to the main body

in the same manner as before. Make sure to securely tighten the servos to the 3D-printed parts.

Step 2: Hand and Leg Assembly

First, assemble the legs using the 3D-printed parts and servo motors, setting the initial positions of the servos accordingly.

Fig 9. Body Assembly

Step 4: Whole Body Structure Assembly

After assembling the hands, legs, and main body, it is time to connect all the parts together. For this, we use custom-made 3D-printed servo horns to ensure smooth movement of the robot.

Step 5: Servo Initial Position

Set the initial servo positions of the robot according to the diagram.

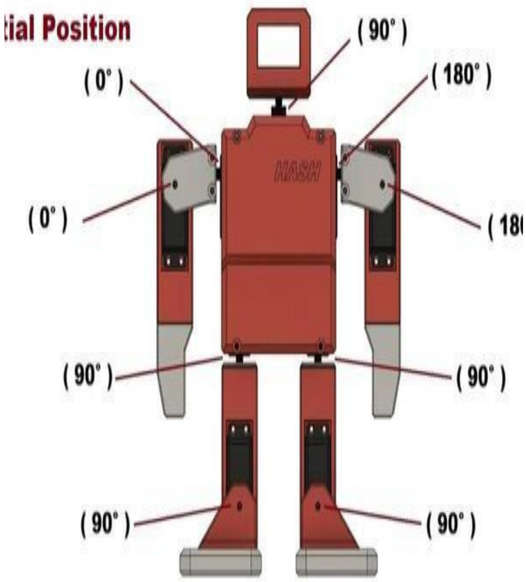

Fig 11. Initial Position

Step 6: Connection Diagram

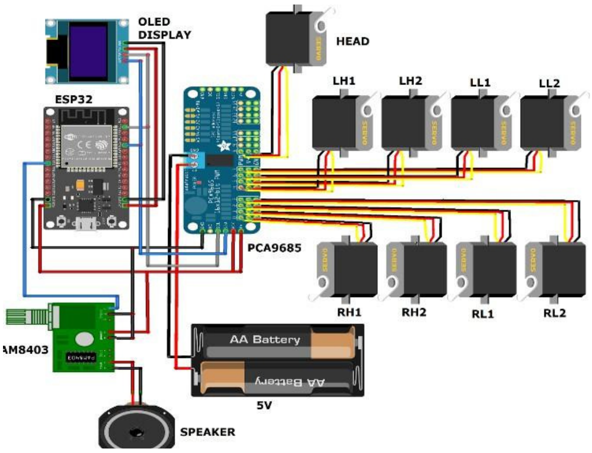

Fig 12. Circuit Diagram

Step 7: Attach the OLED Display to the Head

Now, attach the OLED display to the 3D-printed head and connect it to the ESP32 Dev module.

Step 8: Connect the Servos to the PCA9685 Servo Driver

Then connect all the servo to the servo driver according to the circuit diagram.

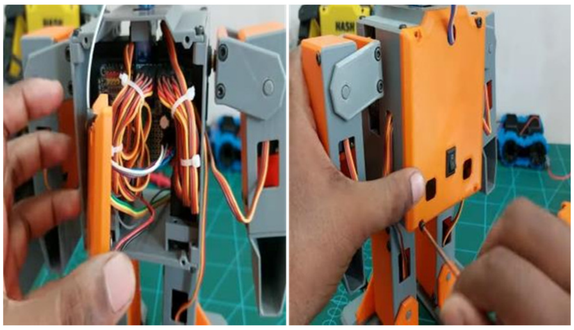

Fig 14. Connecting Servos to the PCA9685 Servo Driver

Step 9: Organize the Cables and Cover the Back

Now for better look organize the cable with some zip ties and cover the back.

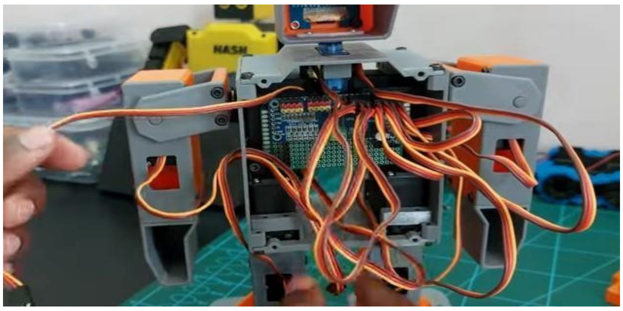

Fig 15. Organize Cables

Step 10: Coding in ESP32

Connect the ESP 32 Dev module using a USB type B cable.

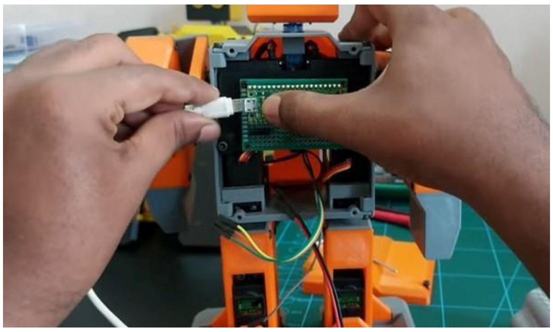

Fig 16. Coding

Step 11: Facial OLED Animations

We need to extract the code for these facial expressions (JPEG or PNG) and integrate them into the coding part.

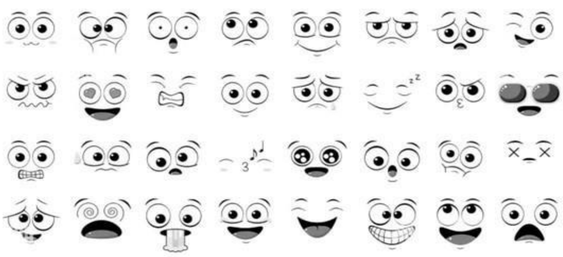

Fig 17. Facial OLED Animations

V. CONCLUSION

The Smart Face Bot successfully integrates speech synthesis, facial expressions, and servo-controlled movements to create an interactive and expressive robotic system. By utilizing the ESP32 microcontroller**,** OLED display, and PCA9685 servo driver, the bot effectively synchronizes voice output with facial expressions. The text-to-speech (TTS) engine enhances communication, while servo-controlled mouth and head movements add realism. The 3D-printed structure provides a customizable and lightweight design. Wireless capabilities enable remote control and future AI integration for improved interactivity. This project demonstrates a cost-effective and scalable solution for human-like robotic interaction. Future enhancements may include emotion recognition and voice command processing for more advanced applications.

VI. REFERENCES

1. Adafruit Industries. (2023). *Adafruit SSD1306 OLED Display Library*. Retrieved from https://learn.adafruit.com/
2. Espressif Systems. (2023). *ESP32 Technical Reference Manual*. Retrieved from https://www.espressif.com/en/products/socs/esp32/documentation
3. Texas Instruments. (2023). *PAM8403 Audio Amplifier Datasheet*. Retrieved from https://www.ti.com/
4. Adafruit Industries. (2023). *PCA9685 16-Channel Servo Driver Guide*. Retrieved from https://learn.adafruit.com/16-channel-pwm-servo-driver
5. Arduino. (2023). *Servo Motor Control using Arduino and ESP32*. Retrieved from https://www.arduino.cc/
6. IEEE Xplore. (2022). *Facial Expression Synthesis for Human-Robot Interaction Using OLED Displays*. Retrieved from https://ieeexplore.ieee.org/
7. OpenAI. (2023). *Text-to-Speech Technology and Its Applications in Robotics*. Retrieved from https://openai.com/research
8. Thingiverse. (2023). *3D Printed Parts for Robotics and Animatronics*. Retrieved from https://www.thingiverse.com/
9. Robotics Today. (2022). *Human-Robot Interaction Using Facial Expressions and Speech Synthesis*. Retrieved from https://www.roboticstoday.com/

ESPRESSIF SYSTEMS. (2023). *ESP32 WI-FI AND BLUETOOTH CAPABILITIES FOR IOT AND ROBOTICS APPLICATIONS*. RETRIEVED FROM HTTPS://WWW.ESPRESSIF.COM/

