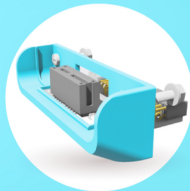


ESP32-S3-BOX-3

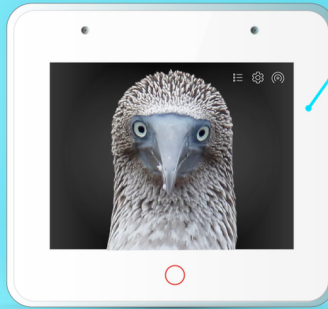
Next-generation AIoT Development Tool



BRACKET



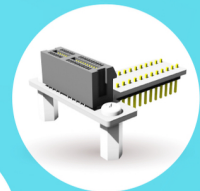
SENSOR



Powered by
ESP32-S3



DOCK



BREAD

ESP32-S3-BOX-3 AIoT Development Kit User Guide

- [中文版本](#)
- [Promotional Video](#)
- [GitHub ESP-BOX SDK](#)

Important Note :

Upon receiving the ESP32-S3-BOX-3, we advise updating the built-in factory firmware to ensure you have the latest bugfixes via Espressif's online firmware flashing tool [ESP Launchpad](#). Up till now, the latest version is **ESP-BOX-3_Demo_V1.2.3**.

Hint - To verify the software version, please go to the 'About Us' Page.

Overview

In a world where innovation combines with artistry, where sharing and collaboration are paramount, we introduce the ESP32-S3-BOX-3, a fully open-source AIoT development kit based on the powerful ESP32-S3 SoC. This next-generation tool brings a positive change to traditional development boards, it features a compact and beautiful enclosure also provides rich assemblies, empowering developers to easily customize and expand its functionality with Espressif trusted software solutions such as ESP-BOX, ESP-SR, ESP-Rainmaker and ESP-Matter. Whether you're prototyping a new project or building a complex IoT system, the ESP32-S3-BOX-3 strikes the perfect balance of form and function, propelling the future of development board ecology.

The ESP32-S3-BOX-3 comes with pre-built firmware supporting offline voice wake-up and speech recognition in both Chinese and English. The ESP-BOX mobile APP facilitates reconfigurable AI voice interaction, enabling you to create personalized commands for controlling smart devices. Furthermore, the firmware showcases various sensor demos and enables users to perform IR learning, transforming the box into a controller for in-house air conditioners.

This guide provides a comprehensive overview of the built-in firmware's capabilities to get you started on your journey. As you read through, you'll gain valuable experience on developing your applications independently. Let's embark on this exciting adventure together!

Target Group

The target group for this product includes individuals and organizations engaged in making, prototyping, and development activities. It is designed to cater to a wide range of users, such as hobbyists, makers, engineers, educators, and professionals seeking versatile and expandable solutions for their projects.

Target Application

- Smart Home
- Smart Agriculture
- Smart City
- Energy Management
- IoT Prototyping and Education
- Industrial Internet of Things
- Smart Retail



ChatGPT AI Chatbot



Library Management System



Stand Reminder



Edge AI in Industry 4.0



Matter Controller



Handheld Gaming Console



DIY Light Strip



Energy Consumption Monitor



Robot Dog

Key Features

- An out-of-the-box entry-level edge AI + HMI application development board.
- FreeRTOS-based four-in-one voice interaction panel: offline voice recognition, network communication, screen display, peripheral control.
- Supports far-field voice interaction with 2 mics.
- Offline voice wake-up and commands recognition with high wake-up rate.
- Allows continuous recognition, wake-up interrupt, and over 200 customizable command words.
- Flexible and versatile home automation solutions: Matter, Home Assistant, ESP-RainMaker.
- Visual drag-and-drop GUI development: LVGL SquareLine Studio, Embedded Wizard, etc.
- Rich development framework: ESP-IDF, Arduino, PlatformIO, Circuit Python, and more.

Hardware Introduction

Contents of the Kit



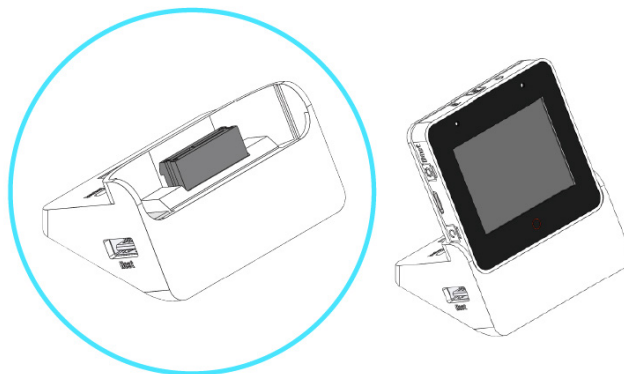
- ESP32-S3-BOX-3, a main unit that can work on its own
- ESP32-S3-BOX-3-DOCK, a functional accessory serving as a stand for the main box
- ESP32-S3-BOX-3-SENSOR, a functional accessory showcasing sensor applications
- ESP32-S3-BOX-3-BRACKET, an adapter accessory for mounting the main box to other devices
- ESP32-S3-BOX-3-BREAD, an adapter accessory facilitating easy connection of the main box to a standard breadboard
- A USB-C Power Cable
- An RGB LED module and Dupont wires for testing

ESP32-S3-BOX-3



ESP32-S3-BOX-3 is the main unit powered by the ESP32-S3-WROOM-1 module, which offers Wi-Fi + Bluetooth 5 (LE) wireless capability as well as AI acceleration capabilities. On top of 512 KB SRAM provided by the ESP32-S3 SoC, the module comes with additional 16 MB Quad flash and 16 MB Octal PSRAM. The board is equipped a 2.4-inch 320 x 240 SPI touch screen (the 'red circle' supports touch), two digital microphones, a speaker, 3-axis Gyroscope, 3-axis Accelerometer, one Type-C port for power and download/debug, a high-density PCIe connector which allows for hardware extensibility, as well as three functional buttons.

ESP32-S3-BOX-3-DOCK



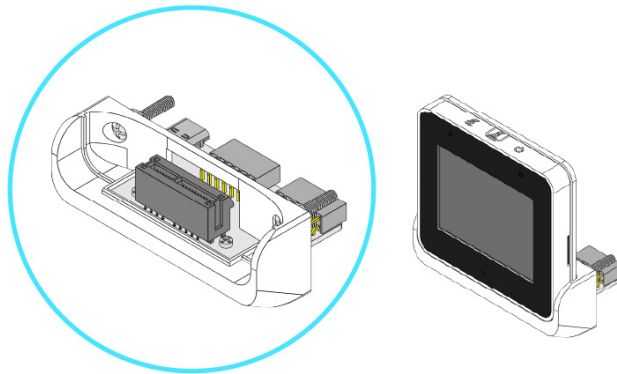
ESP32-S3-BOX-3-DOCK is designed to serve as a stand for ESP32-S3-BOX-3 through the gold fingers, and offers diverse expandability. It has two Pmod™ compatible headers, which allow users to expand additional peripheral modules. These headers offer 16 programmable GPIOs that can supply 3.3 V power to peripherals. One USB Type-A port for connecting to diverse USB devices such as USB camera (with a maximum resolution of 720 P), USB disk, and other HID devices. While the another USB Type-C port for 5 V input power only.

ESP32-S3-BOX-3-SENSOR



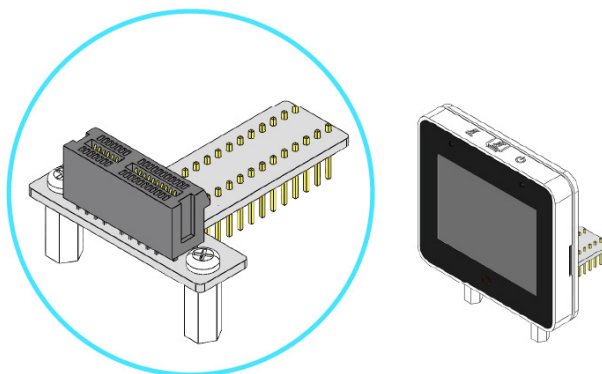
ESP32-S3-BOX-3-SENSOR is a versatile accessory, which integrates Temp&Hum Sensor, IR Emitter and Receiver, Radar Sensor, 18650 Rechargeable Battery slot and MicroSD Card slot. It empowers user to create wide range of innovative projects with ease. Integrate multiple sensors for detection and control, leverage the rechargeable battery for portability, and expand your storage capabilities using the MicroSD Card slot (Expand storage up to 32 GB).

ESP32-S3-BOX-3-BRACKET



ESP32-S3-BOX-3-BRACKET can be utilized to help ESP32-S3-BOX-3 mount to other devices, opens up a multitude of possibilities for transforming non-smart devices into smart ones. Installation of ESP32-S3-BOX-3-BRACKET is straightforward, simply prepare 2 mounting holes and a slot using our template [HERE](#). By leveraging its two Pmod™ compatible headers, user can effectively integrate wireless connectivity, voice control and screen control capabilities. The ESP32-S3-BOX-3-BRACKET adapter empowers you to unleash the full potential of your non-smart devices.

ESP32-S3-BOX-3-BREAD

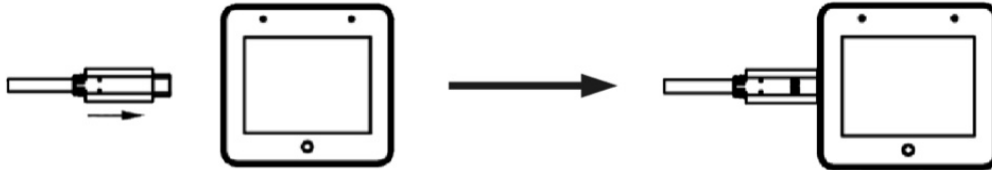


ESP32-S3-BOX-3-BREAD is an adapter that enables easy connection of the ESP32-S3-BOX-3 to a standard breadboard, providing convenience for makers who enjoy hands-on projects to expand and connect with other devices. It utilizes a high-density PCIe connector and two rows of 2.54 mm pitch pins to expose the ESP32-S3's 16 programmable GPIOs, making ESP32-S3-BOX-3-BREAD highly practical and functional.

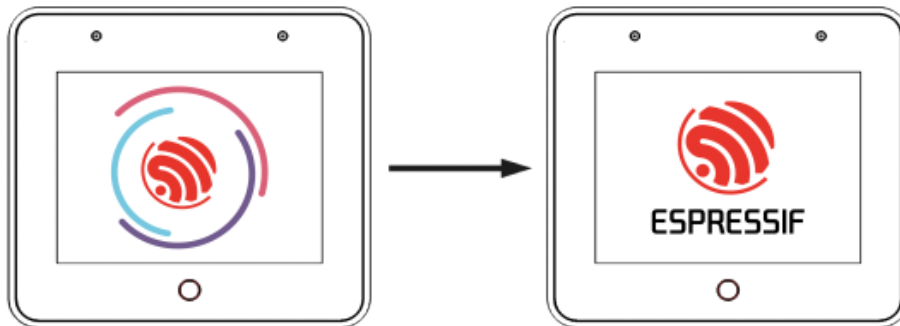
Getting Started

Power on Your Device

1. Power on your device using the USB-C cable.

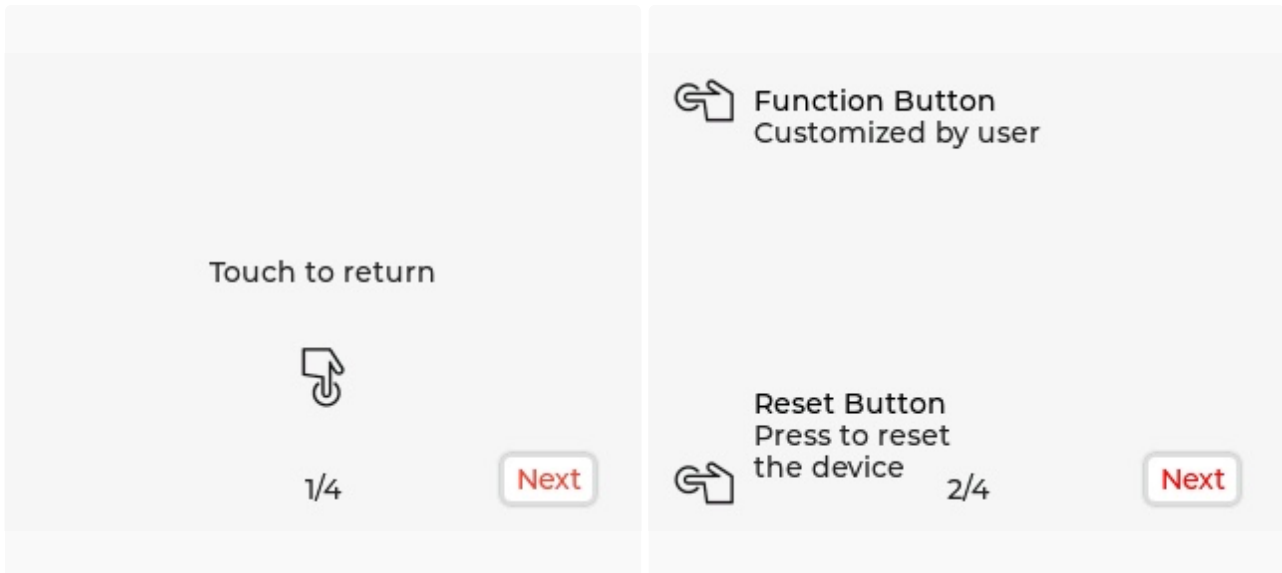


2. Once the device is powered on, the screen will play the Espressif logo boot animation.

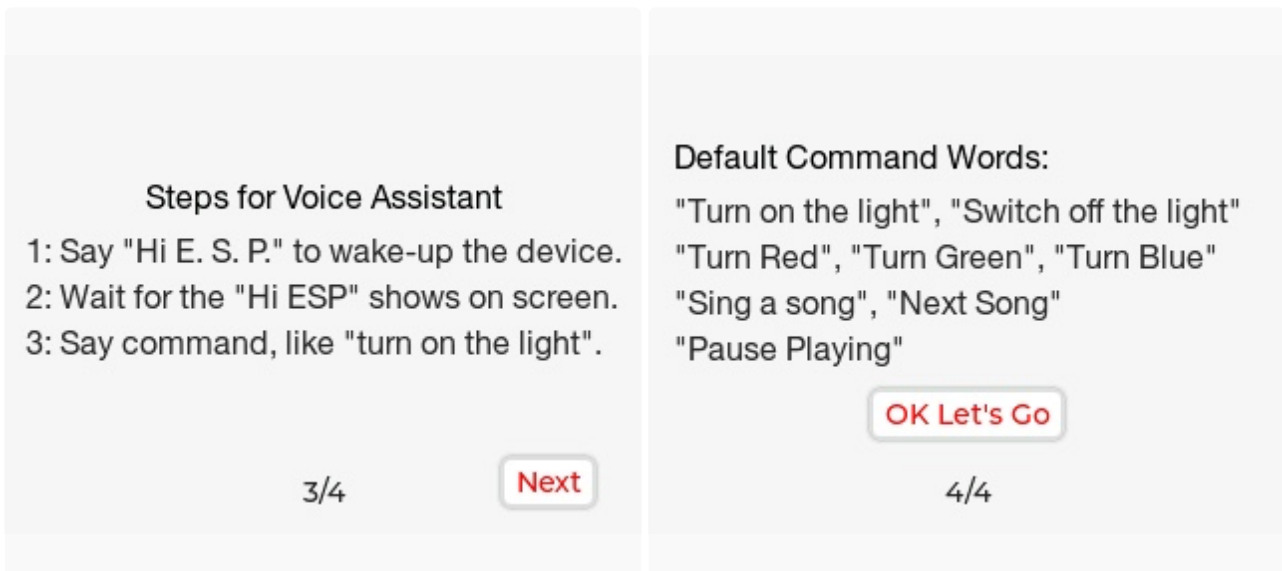


Let's Play Around!

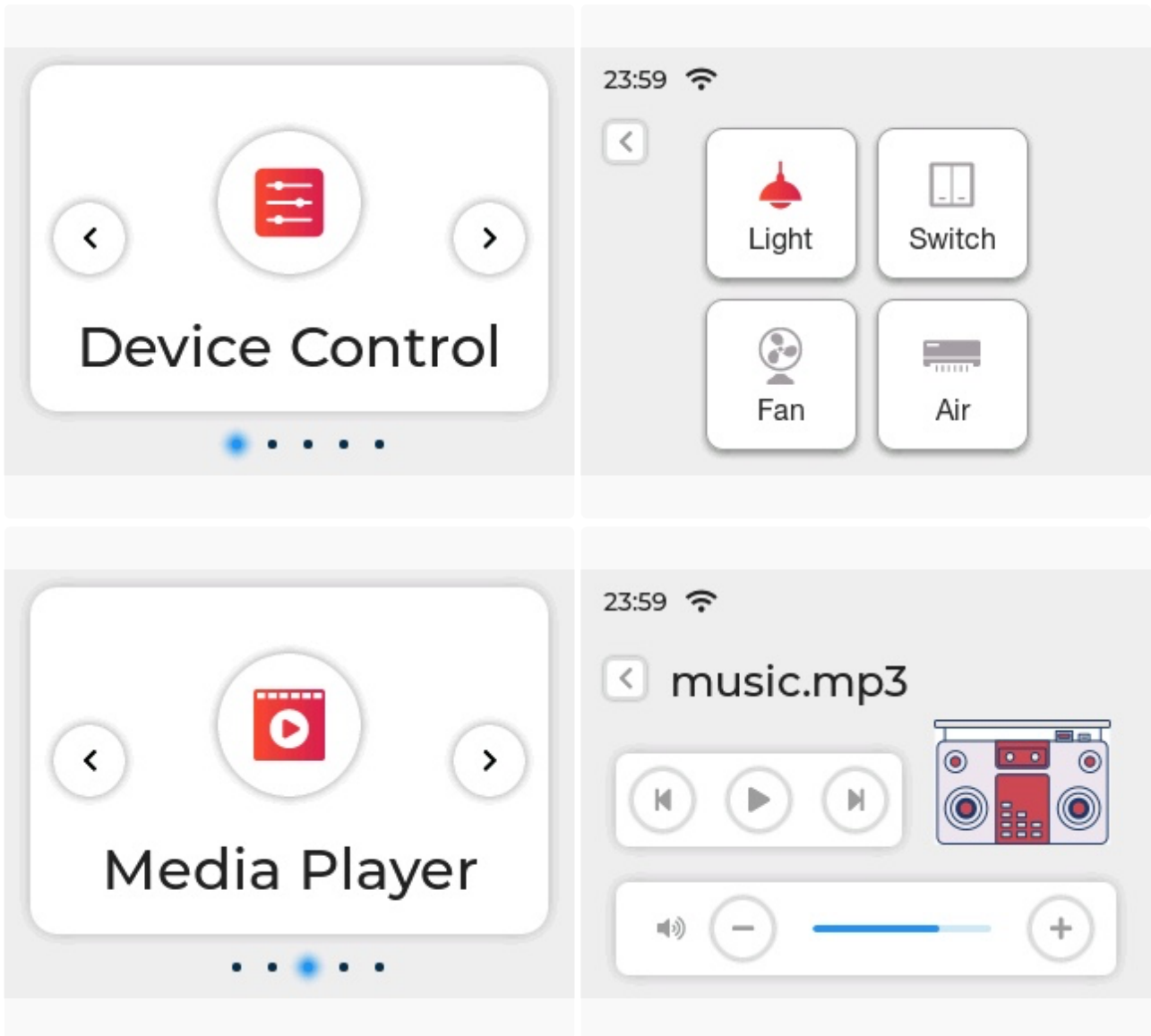
1. The initial two pages of the quick guide provide an overview of the functions of the buttons on your BOX. Press the **Next** button to proceed to the subsequent page.



2. The subsequent pages of the quick guide introduce how to use AI voice control. Tap **OK Let's Go** to access the menu.



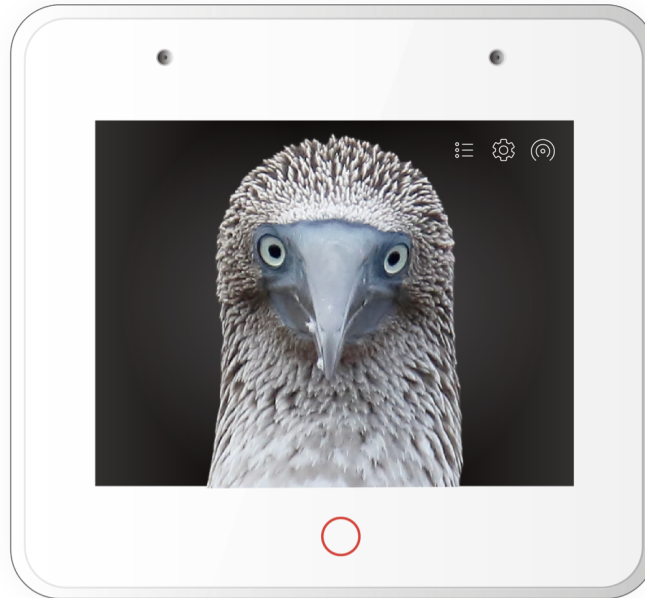
3. The menu consists of six options: **Sensor Monitor**, **Device Control**, **Network**, **Media Player**, **Help**, and **About Us**. You can move between these options by swiping left or right. For instance, you can access the **Device Control** screen, tap on **Light** to toggle the light on or off. Afterward, you can return to the menu, enter the **Media Player** screen, and either play music or adjust the system volume.



4. Press the mute button on the top of the device to disable voice wake-up and speech recognition. Press it again to enable them.

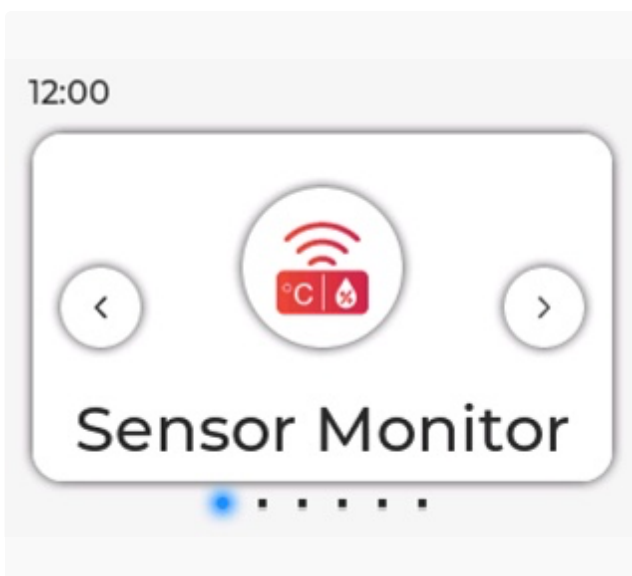


5. Tap the red circle located below the screen to go back to the previous page. This red circle serves as a customizable touch region, allowing users to customize extra touch capabilities according to their preferences.



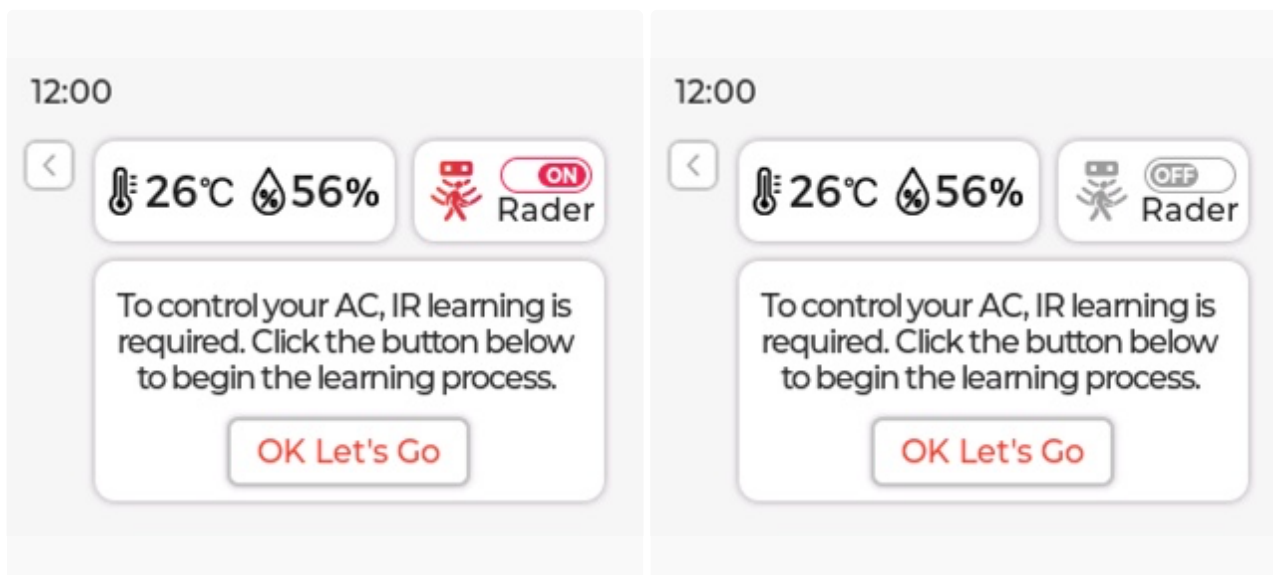
Sensor Monitor

This feature needs ESP32-S3-BOX-3 is mounted to the ESP32-S3-BOX-3-SENSOR dock. ESP32-S3-BOX-3-SENSOR is a versatile accessory, which integrates a Temp&Hum Sensor, a pair of IR Emitter and Receiver, and a Radar Sensor. It empowers user to create sensor network and other sensor applications with ease. The built-in firmware provides a real-time environment Temp&Hum display, a demonstration of human presence monitoring through a 2.4 GHz radar. Additionally, it offers an IR learning interface, allowing users to use the ESP-BOX for IR learning of their air conditioning controller, thereby enabling control over the air conditioner. This learning capability also extends to other in-house IR controllers, such as Fan, TV, projector, providing a truly engaging experience.



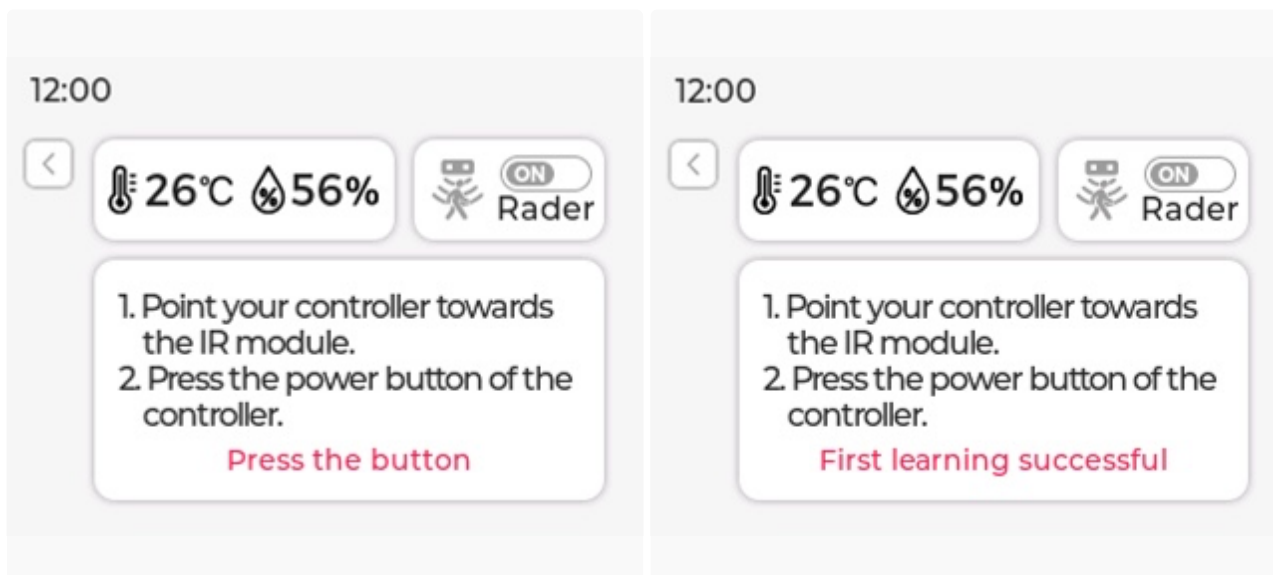
Radar Monitoring

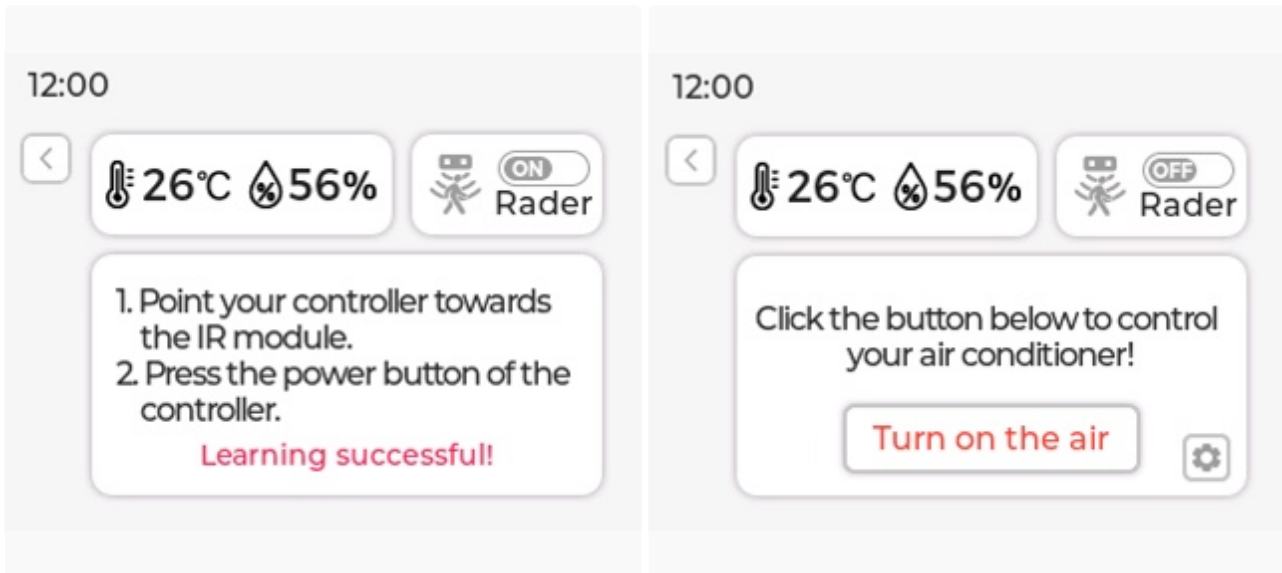
Enter the **Sensor Monitor** screen and tap **ON/OFF** button to enable/ disable radar monitoring feature. When the radar enables switch is set to **ON**, a red body icon will be shown if a person is detected in front of the device, and the icon will turn gray if no person is detected within 2 minutes.



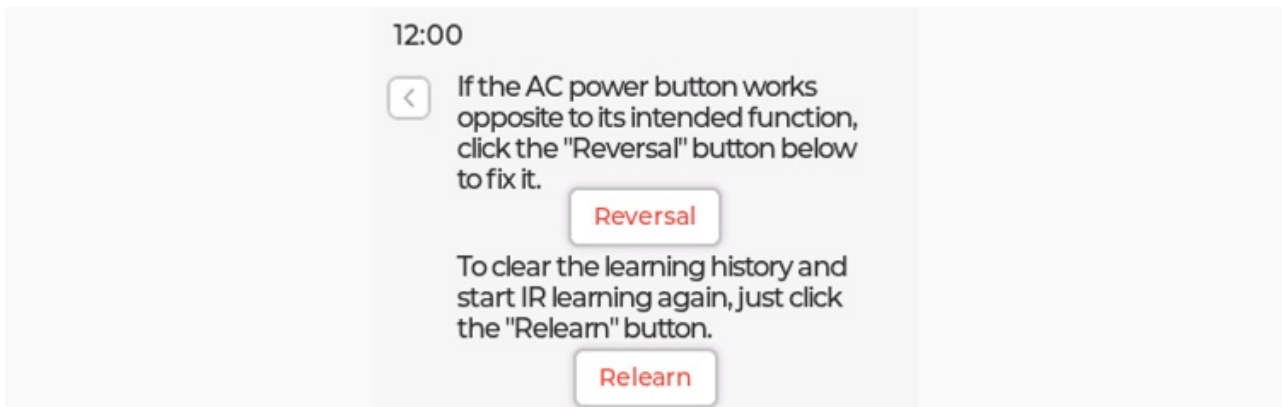
IR Learning

Below the temperature and humidity data and radar functions is the infrared learning module. Currently, this module can only learn the **ON/OFF** function of a remote controller. Please follow the interface instructions to iteratively learn the **ON/OFF** command of your remote controller a total of **four** times. The interface upon successful learning will appear as follows:





Perform an **ON/OFF** test on your air conditioner by pointing the ESP32-S3-BOX-3-SENSOR towards it. If you find that the logic of the air conditioner's **ON/OFF** function is opposite to your expectations, please click the **Reversal** button below to correct it. Of course, you can also click on **Relearn** to learn other remote controllers.

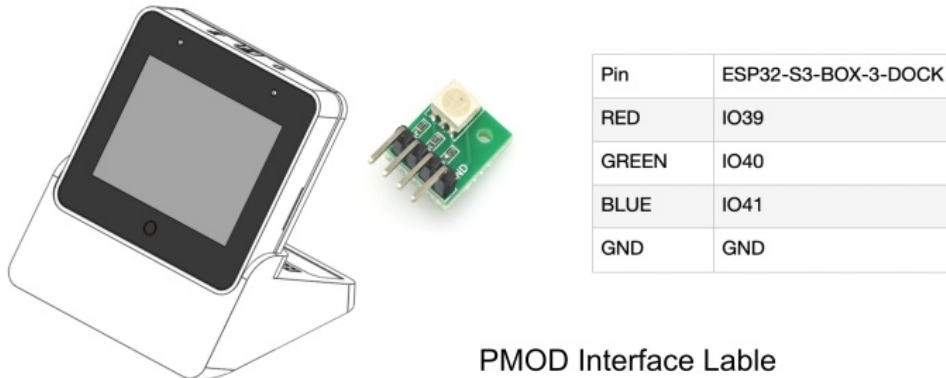


Tips:

- Please be aware that when the ESP32-S3-BOX-3 is not mounted to the ESP32-S3-BOX-3-SENSOR dock, the entire **Sensor Monitor** function will not be operational.
- During the period of experiencing the built-in firmware with ESP32-S3-BOX-3-DOCK, **avoid hot-plugging** ESP32-S3-BOX-3-DOCK and switching to ESP32-S3-BOX-3-SENSOR. This action could result in ESP32-S3-BOX-3-SENSOR **not being recognized and functioning properly**. To resolve this, simply power again ESP32-S3-BOX-3-SENSOR to restore normal operation.
- Due to the power limitations of the infrared emitter and differences among air conditioner brands, the effective control range for the **IR learning** feature has been tested to be between **1 to 1.5 meters**.

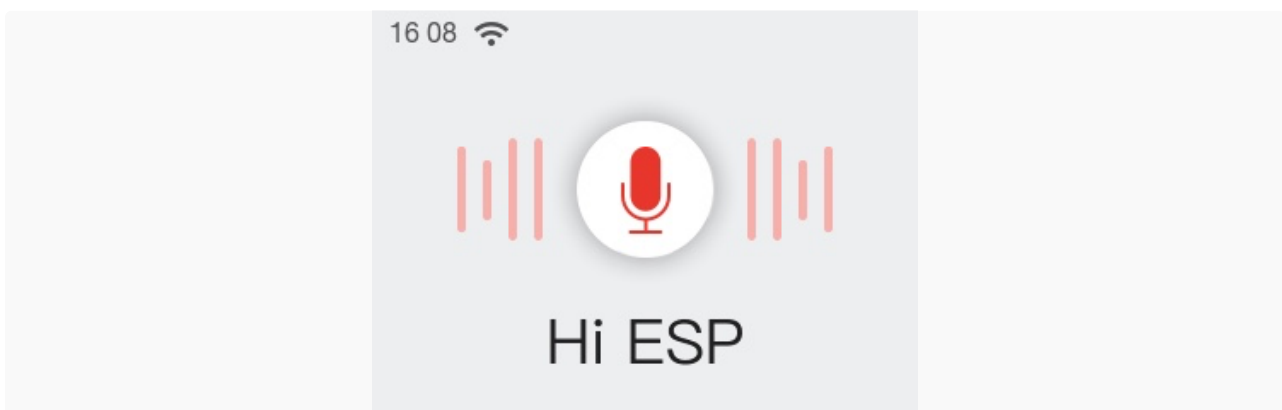
Offline Voice Assistant

Firstly, connect the RGB LED module to your device. Refer to the pin definitions in the table below, and use Dupont wires to connect the R, G, B, and GND pins of the RGB LED module to the female header connector on the Pmod™ interface of the ESP32-S3-BOX-3-DOCK.

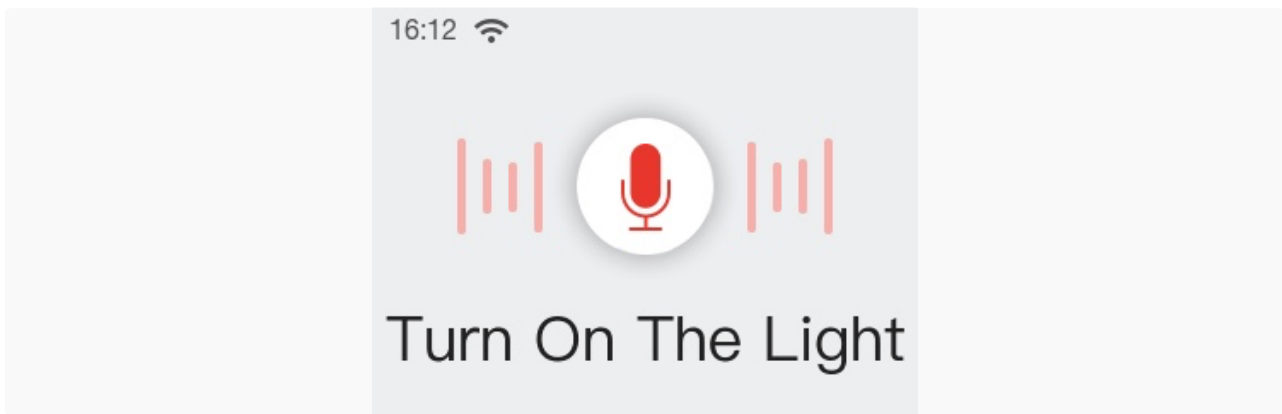


↑	G10	G14	G11	G43/U0TXD	GND	3V3	G21	G19/USB-	G38	G41	GND	3V3
↓	G13	G9	G12	G44/U0RXD	GND	3V3	G42	G20/USB+	G39	G40	GND	3V3

1. You may say "Hi E. S. P." (pronounce in alphabet) at any screen to wake up your device. When it wakes up, the screen will display the wake word you just used. If the wake word isn't displayed, give it another try. The screen below indicates your device is listening.



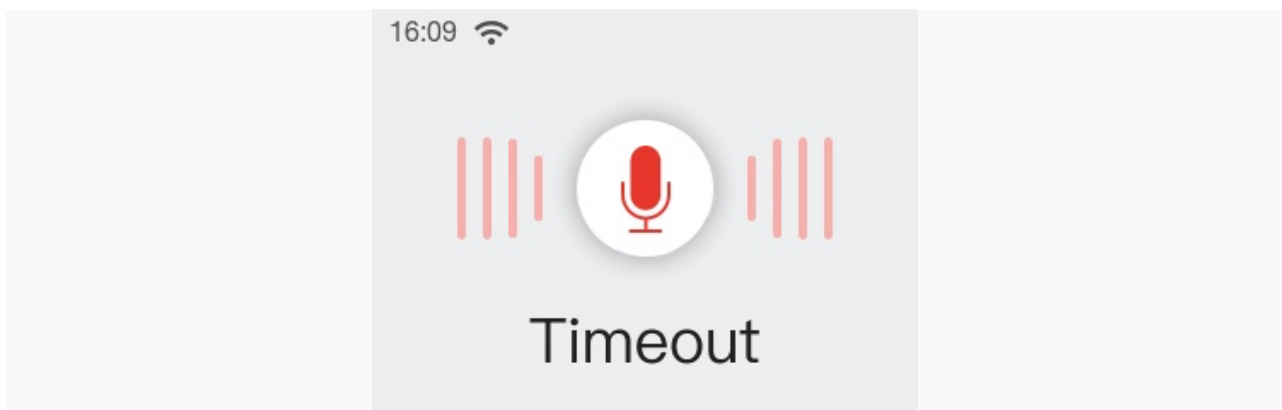
2. Utter a command within 6 seconds after the beep, such as "Turn on the light". You will see the command shown on the screen and the LED light on the module turned on, and you will hear an "OK". If no more commands follow in around 6 seconds, you'll exit the voice control screen.



3. You can use voice commands to enjoy music. First, wake up the device, then say "Sing a song". This will launch the music player and start playing built-in music. You can also use voice commands to pause or skip songs. There are two built-in songs.

Tips:

- If the LED light fails to turn on, check whether the module pins are inserted into the right ports.
- If the BOX doesn't recognize any command within the designated timeframe, a Timeout message will appear, and the screen will exit in approximately 1 second.



4. The default commands are: **turn on the light, turn off the light, turn red, turn green, turn blue, sing a song, next song, pause playing.**

Continuous Speech Recognition

More interestingly, the device supports continuous speech recognition after waking up. This feature adds a natural and seamless quality to voice interaction, enhancing the interactive experience with a human touch.

How to use

- Say "Hi, E. S. P" to wake up the device, and you will hear a beep.
- Speak your command. Upon successful recognition, you will hear "OK", and the device will remain ready for further commands.
- If no command is identified, the device will wait. Should no commands be given within 6 seconds, the device will autonomously exit the voice control screen and you need to wake it up again.

Attention

- If your commands are repeatedly unrecognized by the device, allow it to time out and then wake it up again before utilizing the feature.
- Upon saying the wake-up word, please do not move the device, as this could impact command recognition.
- We recommend that voice commands comprise 3-5 words for optimal results.
- Currently, the device cannot recognize commands while it's playing beep.

Voice Command Customization

The ESP32-S3-BOX-3 is also equipped with Espressif proprietary AI Speech Recognition System, which allows you to customize commands through our ESP BOX APP. We will take the LED light on the module as an example, to show how to create your own voice commands. For a deeper understanding of the algorithms involved, please refer to the introduction of the ESP-SR speech model in the Technical Architecture.

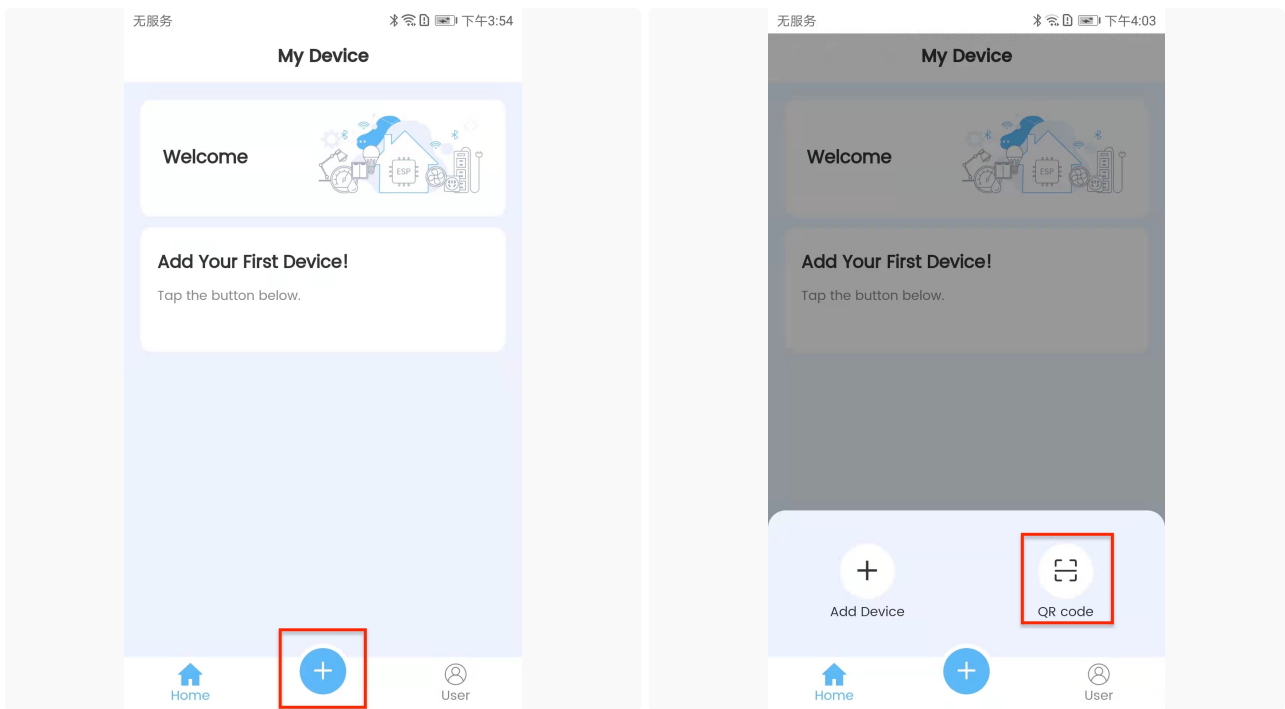
1. Connect to the ESP BOX mobile app

1.1. Enter `Network`, and tap `To install APP` at the upper right corner. Scan the QR Code or search "ESP BOX" in App Store or Google Play to install the app.

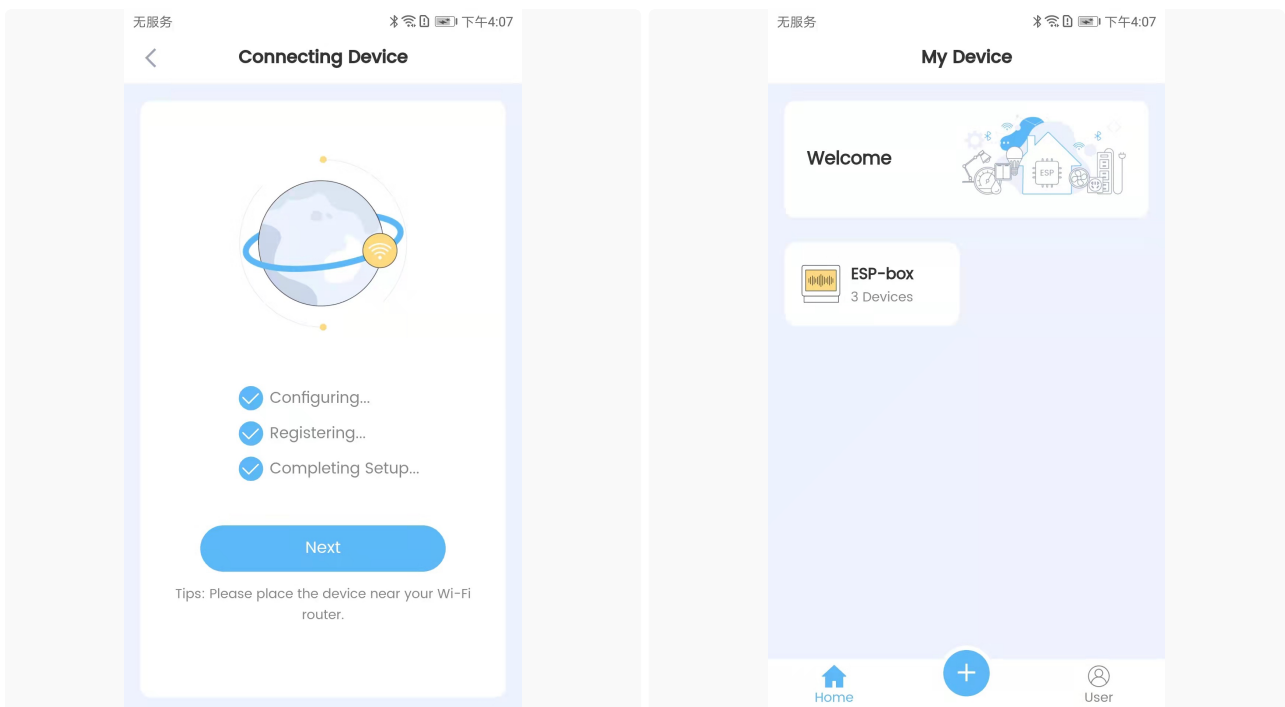


1.2. If you are new to this app, please register an account first.

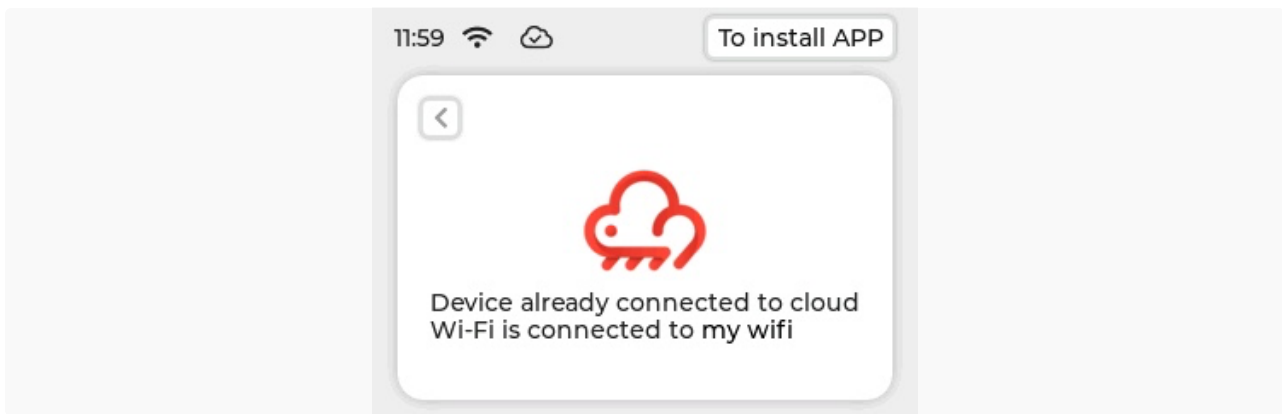
1.3. Sign in with your ESP BOX account and turn on the Bluetooth on your phone. Tap `+` at the bottom of the screen, and scan the QR code on your device to set up the network.



1.4 After adding the device, you will see the following prompts:

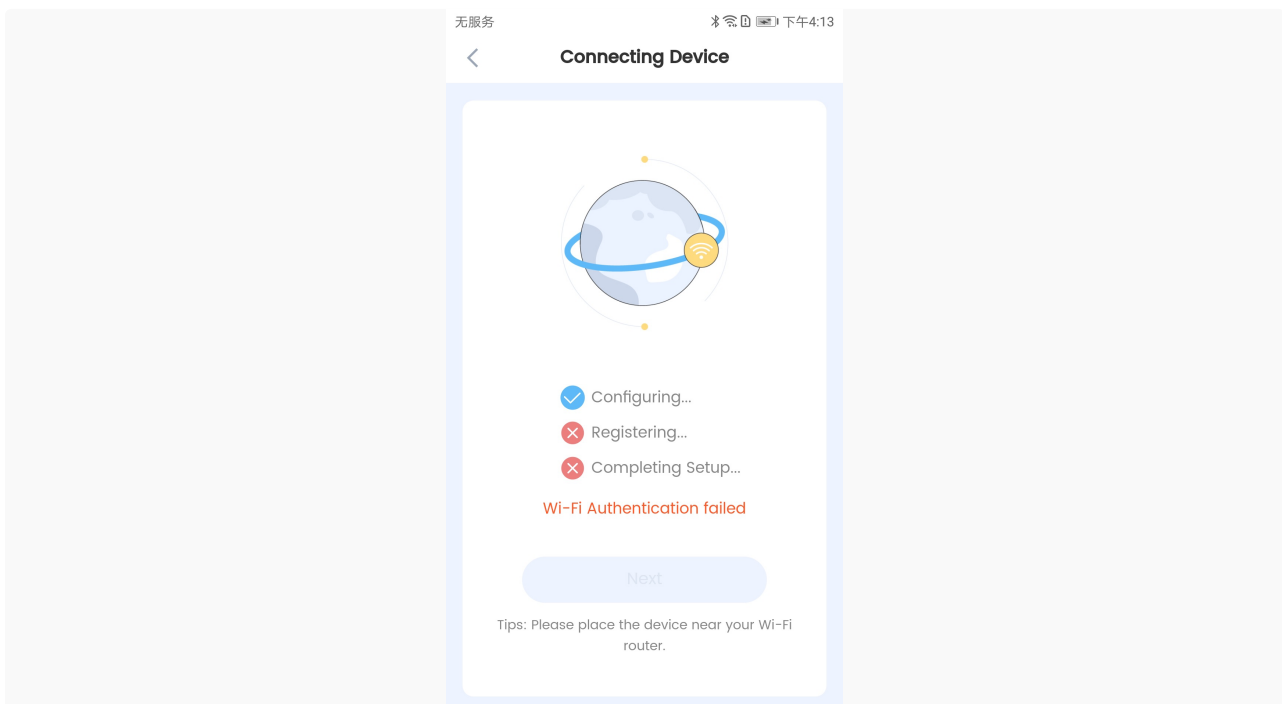


And this prompts in the box:



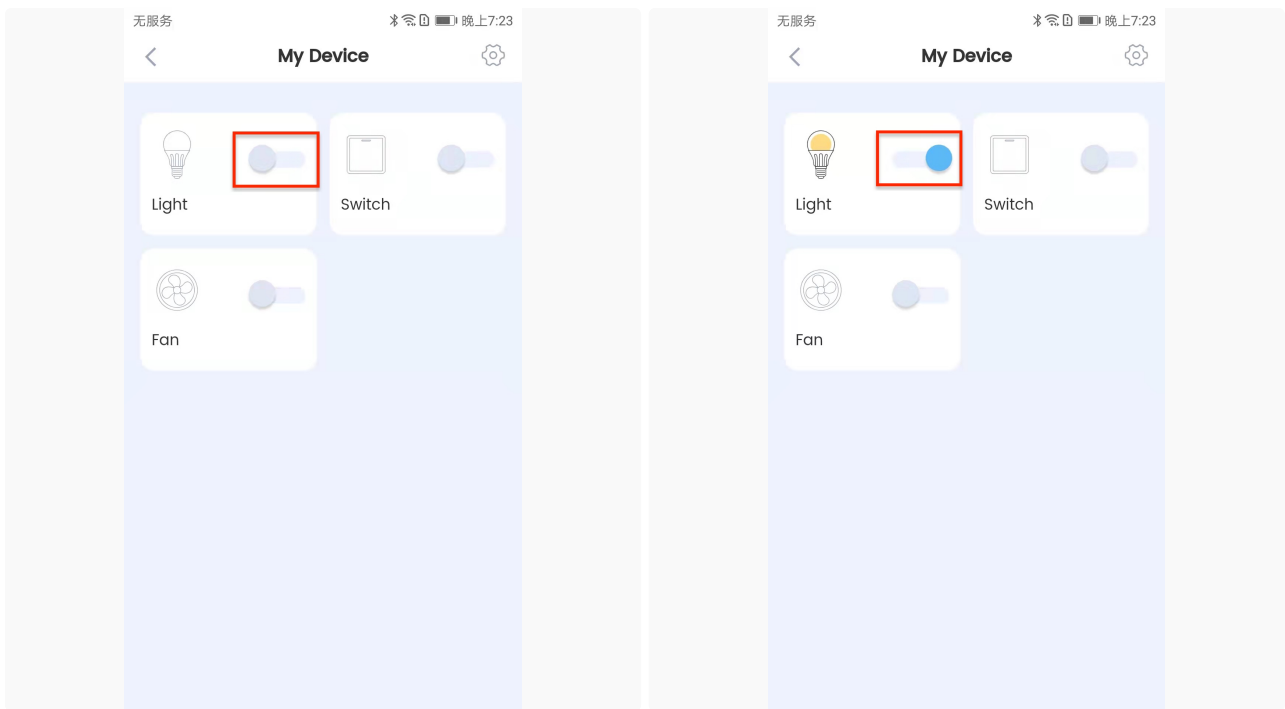
Tips:

- Please do not exit the QR code page during the network provisioning.
- Make sure you connect the device to 2.4 GHz Wi-Fi instead of 5 GHz, and enter the correct Wi-Fi password. If the Wi-Fi password is incorrect, the prompt "Wi-Fi Authentication failed" will pop up.
- Long press the **Boot** button (i.e. **Function** button) for 5 seconds to clear the network information and restore the device to factory settings. After the device is reset, if the QR code or Bluetooth is not working, please restart your device by pressing the **Reset** button.

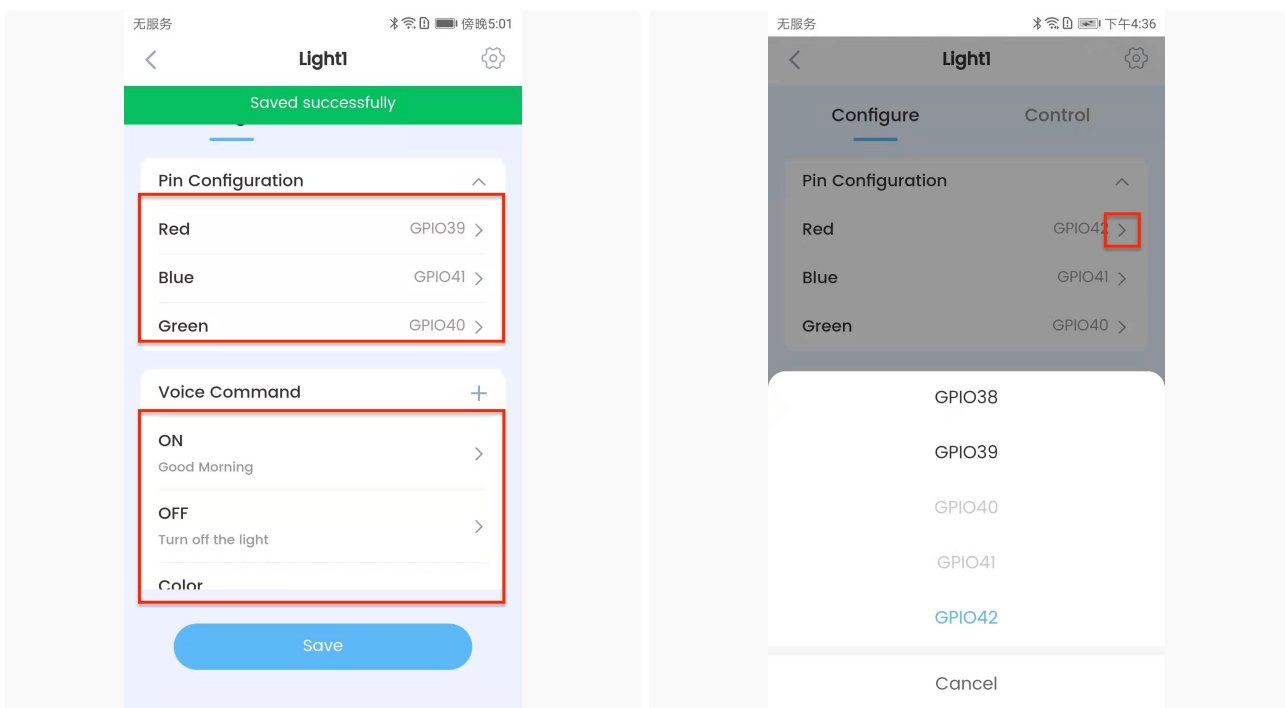


2. Customize Voice Commands

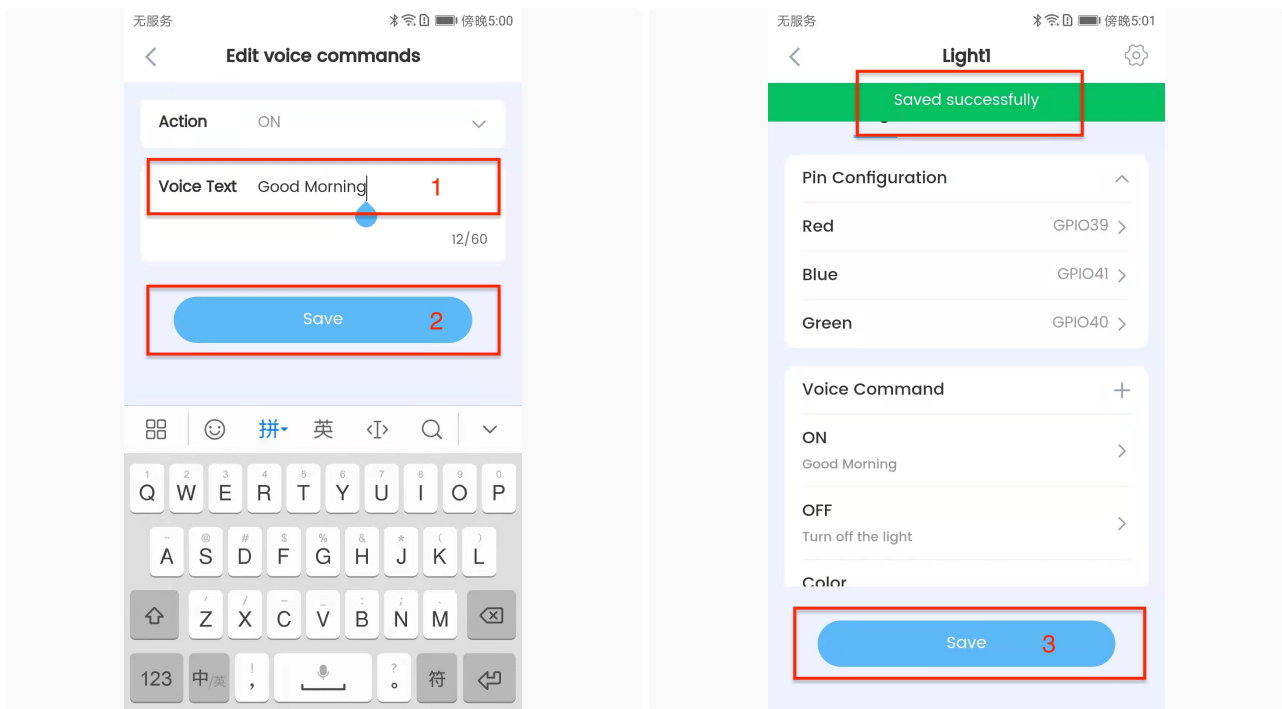
2.1. Select the ESP-BOX device icon and access the screen displayed below. You can conveniently turn the light on or off by simply toggling the button, as indicated in the image. You may need to develop the Fan and Switch functionalities by yourself.



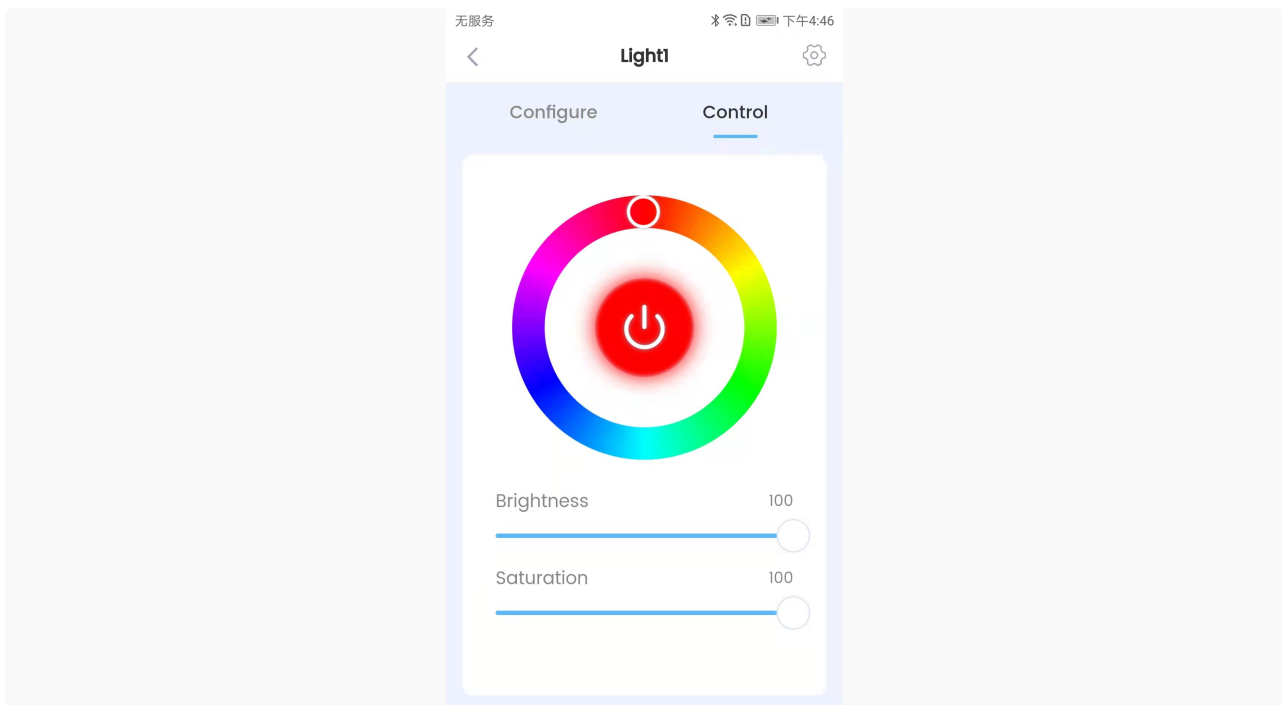
2.2 Tap **Light**, and the **Configure** tab shows the default pin information and commands. The pins for Red, Green, and Blue can be changed as needed.



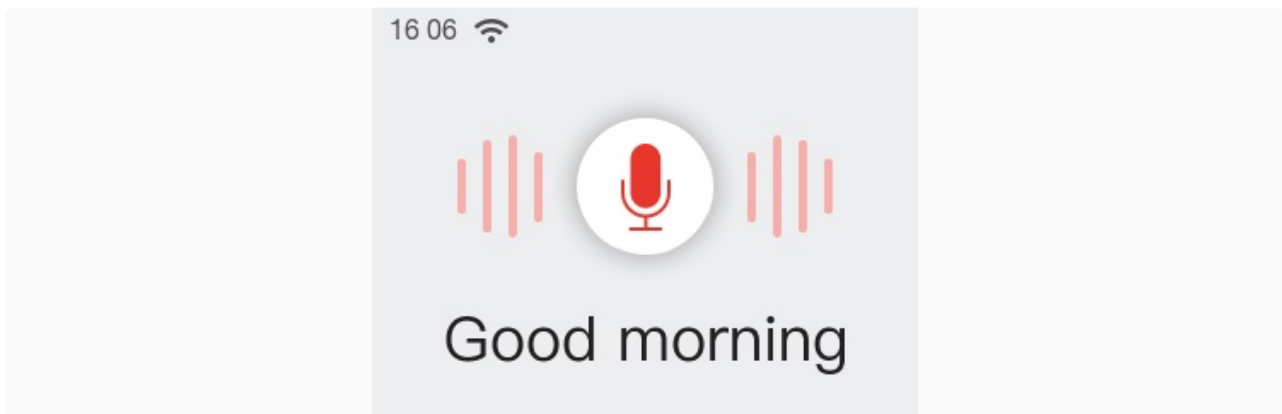
2.3 In the **Configure** tab, you may also customize commands to turn on or off the light and change its color. For example, you can set "Good Morning" as the command to turn on the light. Click **Save** to return to the previous screen. Then click **Save** again as shown below.



2.4 In the **Control** tab, you may adjust the color, brightness, and saturation of the light.



2.5 Now, you can try out your new command! First, say "Hi E. S. P." to wake up your device. Then say "Good Morning" within 6 seconds to turn on the light. The new command will show on the screen with the module light turned on.



For effective command performance, kindly consider the following:

- **Command Length:** A command should consist of 2-8 words. While crafting a sequence of commands, it's advisable to maintain a consistent length across them.
- **Avoid Repetition:** Please do not include shorter commands within longer ones. For example, if you create both "turn on" and "turn on the light" commands, the shorter "turn on" command won't be identified.

3. Development Guide for mobile APP

The ESP-BOX APP is built based on the commercial software architecture, Nova Home APP, which provides fancy UI and better user experience but has limited openness, customers need commercially contact us to access it by email sales@espressif.com.

For developer, we offer ESP RainMaker solution, which includes publicly available development document and API guide. Below is the pathway to access them:

- [ESP RainMaker Get Started](#)
- [RainMaker App APIs Definitions](#)
- [ESP RainMaker Programming Guide](#)

Switch Voice Model Language

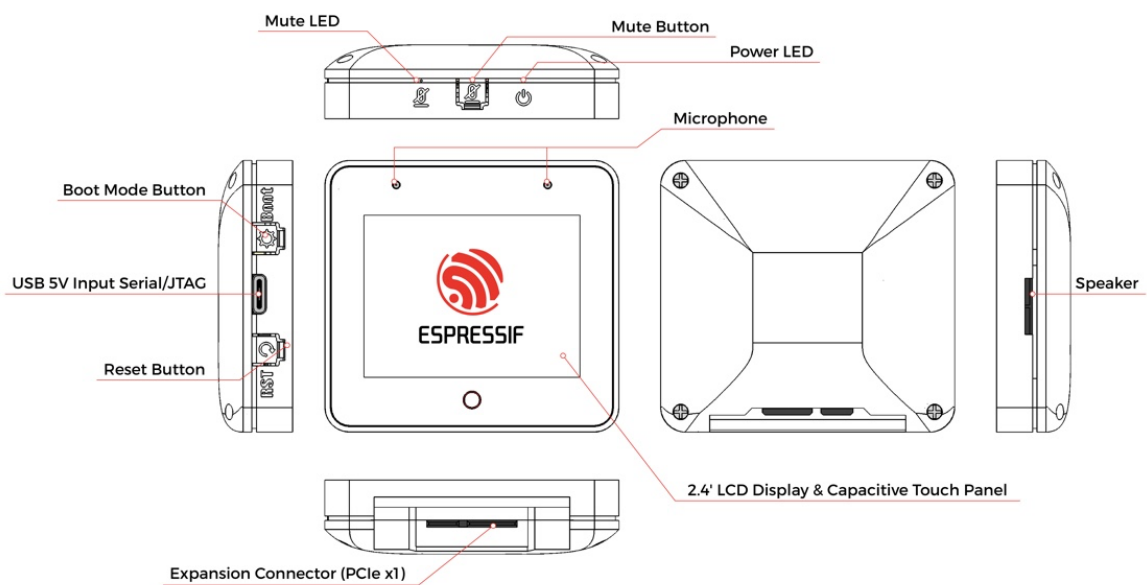
Access the **About Us** details page. In this page, press **Boot** button, the device will enter **factory Mode** voice model language setting page. Here, you can personally select the voice model. After switching the model, the device will reset and restart automatically.



Hardware Overview

ESP32-S3-BOX-3

Orthographic View

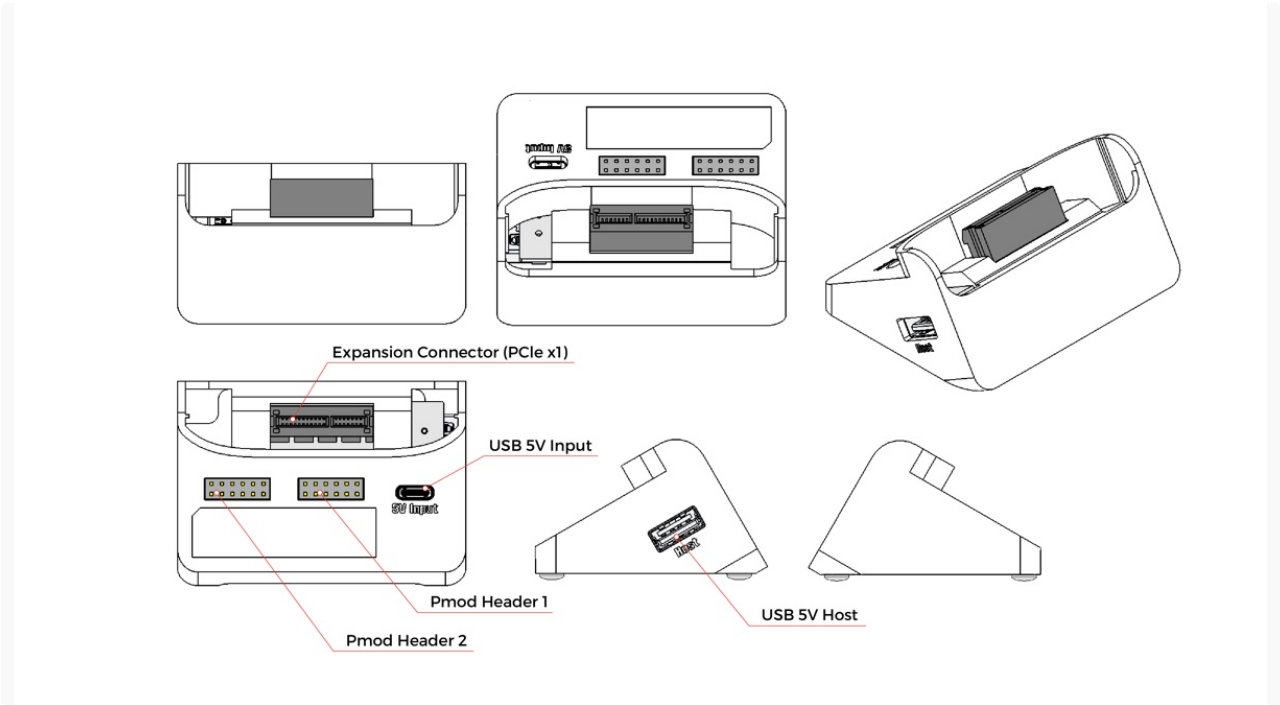


Technical Specification

Microcontroller			
Type:	ESP32-S3	CPU:	Dual-Core Xtensa® 32bit LX7 up to 240 MHz
Memory			
SRAM:	512 KB	ROM:	384 KB
PSRAM:	Octal SPI, 16 MB	PSRAM Speed:	120 MHz (Experimental Feature)
External Flash:	Quad SPI, 16 MB		
AI Feature			
AI Algorithm	Neural Network, Acoustic algorithm, etc.	Computing Acceleration	Vector, Complex number, FFT, etc.
Wireless			
Wi-Fi:	2.4 GHz, IEEE 802.11b/g/n	Bluetooth® LE:	Bluetooth® 5 LE and Bluetooth® mesh
Display			
Display Type:	2.4-inch LCD	Display Resolution:	240 x 320 pixels
Display Interface:	SPI	Interface Speed:	40 MHz
Touch Type:	Capacitive	Touch Points:	10
Driver IC	ILI9342C		
Audio Input			
Microphone Type:	Dual Mic	ADC Model:	ES7210
Mute:	Supported		
Audio Output			
Speaker Model:	8 Ohm 1 W	PA Model:	NS4150
Codec Model:	ES8311		
Sensor			
Sensor Type:	3-axis Gyroscope, 3-axis Accelerometer	Sensor Model:	ICM-42607-P
Interface			
Type:	USB Type-C	Usage:	Power, USB download/JTAG debug, general USB device functions
Type:	Goldfinger	Usage:	I/O Expansion
User Interface			
Onboard Buttons	Reset, Boot, Mute	Onboard LEDs	Power LED, Mute LED
OS / SDK			
Original OS:	FreeRTOS	SDK:	ESP-IDF
Outline			
Dimensions:	61 x 66 x 16.6 mm	Weight:	292 g
Power			
USB-C Power:	5 V - 2.0 A Input	Battery:	N/A

ESP32-S3-BOX-3-DOCK

Orthographic View



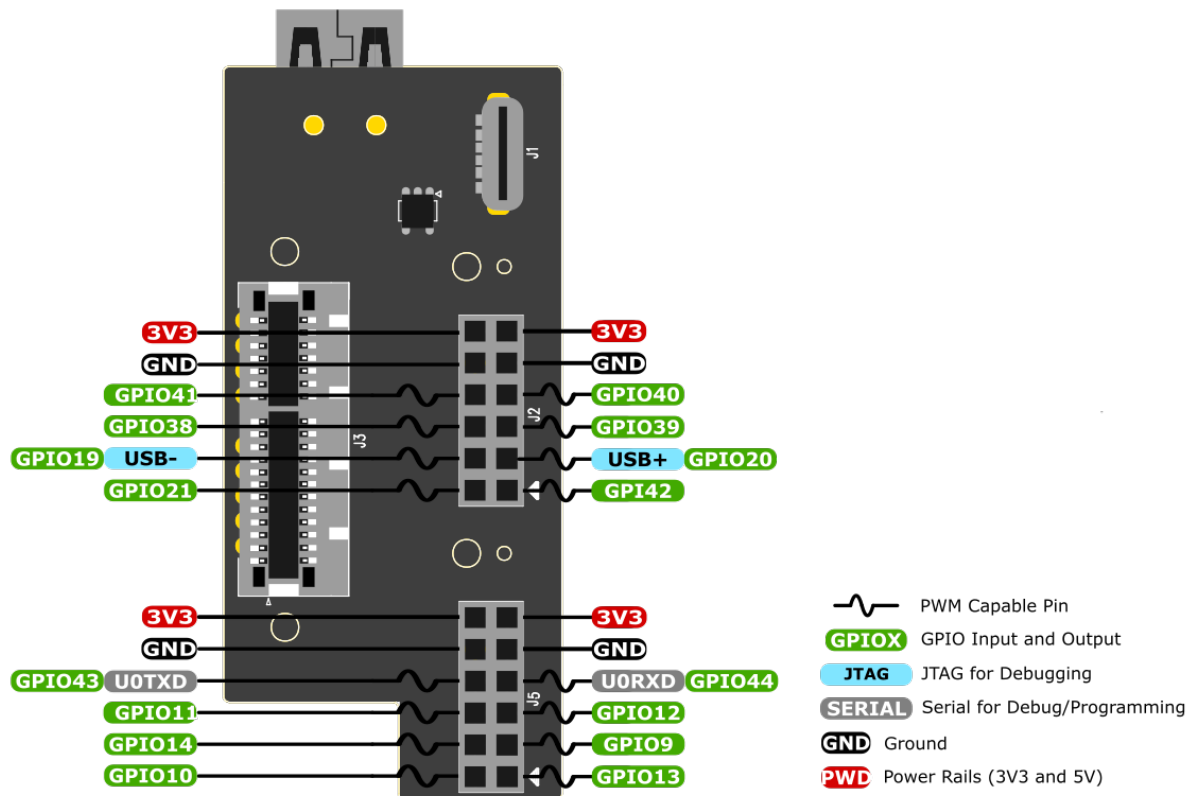
Technical Specification

ESP32-S3-BOX-3-DOCK			
Type	Number	Details	Usage
12-pin Female Header	2	8 I/O (Pmod™ Compatible), 3.3 V Power Output	GPIO, I2C, SPI, UART, RMT, LEDC, etc.
USB Type-A	1	5 V Power Output, USB Host	Connect to diverse USB devices such as USB camera, USB disk, and other HID devices
USB Type-C	1	5 V Power Input	5 V power input only
PCIe Connector	1	36 Pin, 1.00 mm (.0394") pitch, Accepts .062" (1.60 mm) card	Vertical mounting goldfinger

- [Digilent Pmod™ Interface Specification](#)

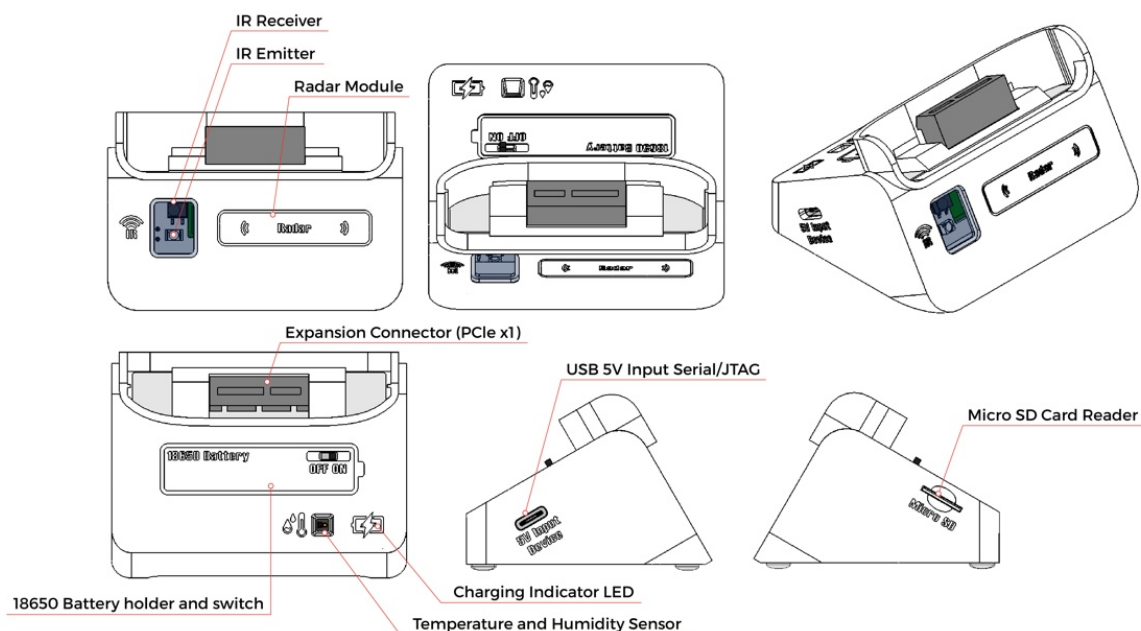
Pinout Diagram

ESP32-S3-BOX-3-DOCK



ESP32-S3-BOX-3-SENSOR

Orthographic View



Technical Specification

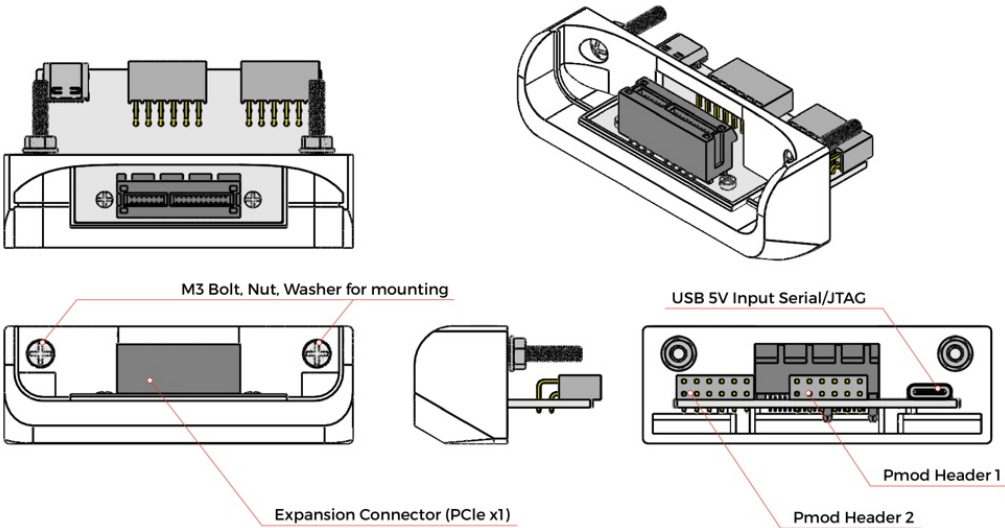
ESP32-S3-BOX-3-SENSOR			
Type	Number	Details	Usage
Radar Sensor (MS58-3909S68U4)	1	Frequency band: 5.8 GHz; Ultimate Power Consumption: 40 uA; Sensing Range: Around 2 meters	Human presence detection
Infrared Sensor	2	A pair of IR emitter (IRM-H638T) and receiver (IR67-21C/TR8) tubes around 4 meters	Infrared control applications
Temp&Hum Sensor (AHT30)	1	Temp Range: -40 ~ + 120 °C/± 0.5 °C; Hum Range: 0 ~ 100% RH/± 3% RH (25 °C)	Temperature and humidity monitoring
External Storage	1	MicroSD card slot (Maximum 32GB)	Expand storage capabilities
Battery	1	18650 Rechargeable battery slot	Sustainable and portable power supply
Switch	1	2-speed toggle switch	Control of battery charging and discharging to protect the 18650 battery
Indicator LED	1	Charging indicator LED	Red: Charging; Green: Fully charged
USB Type-C	1	5 V Power Input, USB Device	Power, USB download/JTAG debug, general USB device functions
PCIe Connector	1	36 Pin, 1.00 mm (.0394") pitch, Accepts .062" (1.60 mm) card	Vertical mounting goldfinger

Tips:

- When charging the 18650 battery, please switch the toggle to 'ON' and plug the ESP32-S3-BOX-3 into the ESP32-S3-BOX-3-SENSOR.
- Again, during the period of experiencing the built-in firmware with ESP32-S3-BOX-3-DOCK, **avoid hot-plugging** ESP32-S3-BOX-3-DOCK and **switching to ESP32-S3-BOX-3-SENSOR**. This action could result in ESP32-S3-BOX-3-SENSOR **not being recognized and functioning properly**. To resolve this, simply power again ESP32-S3-BOX-3-SENSOR to restore normal operation.

ESP32-S3-BOX-3-BRACKET

Orthographic View

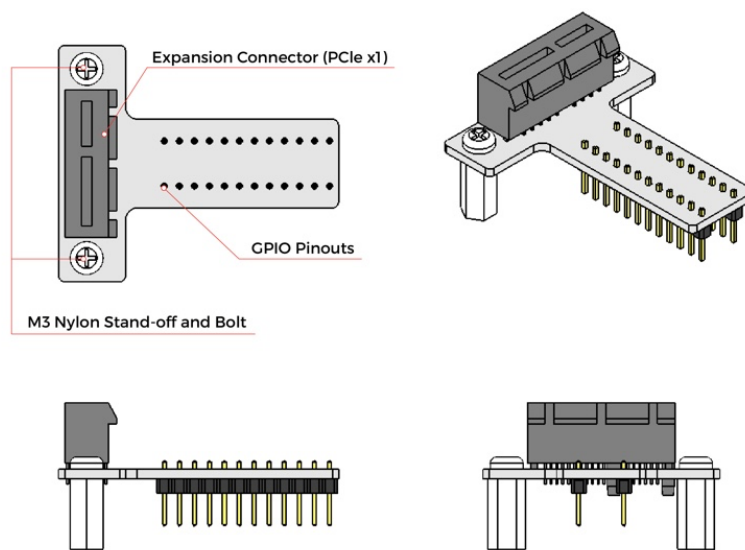


Technical Specification

ESP32-S3-BOX-3-BRACKET			
Type	Number	Details	Usage
12-pin Female Header	2	8 I/O (Pmod™ Compatible), 3.3 V Power Output	GPIO, I2C, SPI, UART, RMT, LEDC, etc.
USB Type-C	1	5 V Power Input, USB Device	Power, USB download/JTAG debug, general USB device functions
PCIe Connector	1	36 Pin, 1.00 mm (.0394") pitch, Accepts .062" (1.60 mm) card	Vertical mounting goldfinger
Bolt	2	M3 Bolt, Nut, and Washer	Mount materials and fasten components together

ESP32-S3-BOX-3-BREAD

Orthographic View

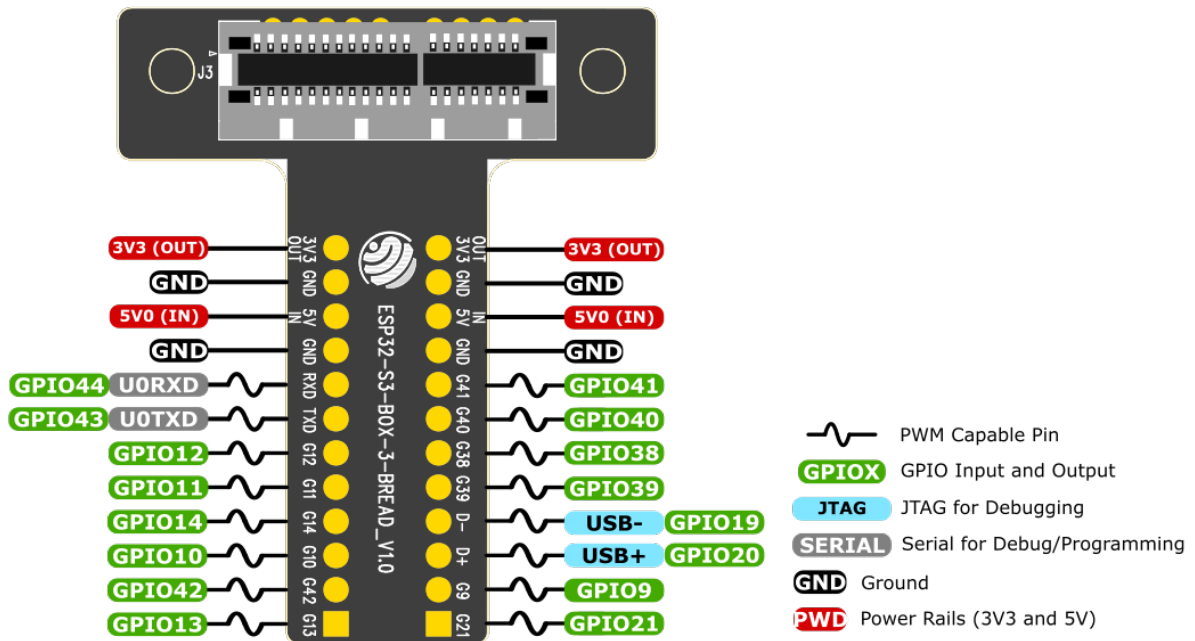


Technical Specification

ESP32-S3-BOX-3-BREAD			
Type	Number	Details	Usage
12-pin Male Header	2	8 I/O, 3.3 V Power Output, 5 V Power Input	GPIO, I2C, SPI, UART, RMT, LEDC, etc.
USB Type-C	1	5 V Power Input, USB Device	Power, USB download/JTAG debug, general USB device functions
PCIe Connector	1	36 Pin, 1.00 mm (.0394") pitch, Accepts .062" (1.60 mm) card	Vertical mounting goldfinger

Pinout Diagram

ESP32-S3-BOX-3-BREAD



FCC Regulations

This device complies with part 15 of the FCC Rules. Operation is subject to the following two conditions:

(1) This device may not cause harmful interference, and (2) this device must accept any interference received, including interference that may cause undesired operation. This device has been tested and found to comply with the limits for a Class B digital device, pursuant to Part 15 of the FCC Rules. These limits are designed to provide reasonable protection against harmful interference in a residential installation. This equipment generates, uses, and can radiate radio frequency energy and, if not installed and used in accordance with the instructions, may cause harmful interference to radio communications. However, there is no guarantee that interference will not occur in a particular installation. If this equipment does cause harmful interference to radio or television reception, which can be determined by turning the equipment off and on, the user is encouraged to try to correct the interference by one or more of the following measures:

- Reorient or relocate the receiving antenna.
- Increase the separation between the equipment and receiver.
- Connect the equipment to an outlet on a circuit different from that to which the receiver is connected.
- Consult the dealer or an experienced radio/TV technician for help.

FCC Note Caution: Changes or modifications not expressly approved by the party responsible for compliance could void the user's authority to operate the equipment.

RF Exposure Information

This device meets the government's requirements for exposure to radio waves.

This device is designed and manufactured not to exceed the emission limits for exposure to radio frequency (RF) energy set by the Federal Communications Commission of the U.S. Government.

This device complies with FCC radiation exposure limits set forth for an uncontrolled environment. In order to avoid the possibility of exceeding the FCC radio frequency exposure limits, human proximity to the antenna shall not be less than 20 cm during normal operation.

The guide only gives you a brief idea of how to use the latest firmware on your BOX series of development boards. Now, you may start writing programs, and embark on your IoT journey!