Infonique Sound Recognition

|  |  |  |
| --- | --- | --- |
| Prepared by | Date | Version |
| Bing Ran | 03/20/2025 | 1.0 |

# 1 Abstract

This document details the implementation of a sound recognition system using the ESP32, covering both hardware and software aspects. The system is designed to recognize voice commands, process them, and respond with spoken feedback using pre-recorded audio files. The hardware includes an ESP32 Dev module, a VC-02 voice recognition module, an I2S audio output system, and an OLED display for visual feedback. The software integrates voice detection, serial communication, and audio playback . This document provides an in-depth exploration of system design, firmware development, and practical considerations, offering a comprehensive guide for developers interested in embedded sound recognition applications.

# Document History

|  |  |  |  |
| --- | --- | --- | --- |
| **Date** | **Rev** | **Modifier** | **Changes** |
| **20-March-2025** | 1.0 | Bing Ran | First Draft |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |

# Contents

[Abstract 2](#__RefHeading___Toc1131_1720759766)

[Document History 2](#__RefHeading___Toc1133_1720759766)

[Contents 3](#__RefHeading___Toc1135_1720759766)

[Table of Figures 4](#__RefHeading___Toc1137_1720759766)

[Table of table 4](#__RefHeading___Toc1139_1720759766)

[1 Introduction 5](#__RefHeading___Toc1141_1720759766)

[2 Drone main board 5](#__RefHeading___Toc1143_1720759766)

[2.1 Drone Main Board Pinout 6](#__RefHeading___Toc1145_1720759766)

[2.2 Connectors 7](#__RefHeading___Toc1093_2042270136)

[2.2.1 Power input connector 7](#__RefHeading___Toc1095_2042270136)

[2.2.2 Motor 7](#__RefHeading___Toc1097_2042270136)

[2.2.3 Motor control connectors 8](#__RefHeading___Toc1099_2042270136)

[2.2.4 SD card connector 8](#__RefHeading___Toc1101_2042270136)

[2.2.5 Barometer BME/BMP 280 module 9](#__RefHeading___Toc1147_1720759766)

[2.2.6 Accelerator & gyro sensor MPU6050 module 9](#__RefHeading___Toc1105_2042270136)

[2.2.6 GPS GT-U7 module 10](#__RefHeading___Toc1151_1720759766)

[GT-U7 module able to communicate through UART communication 10](#__RefHeading___Toc1109_2042270136)

[2.2.7 Radio frequency nRF24L01+ 2.4GHz wireless RF transceiver module 10](#__RefHeading___Toc1111_2042270136)

[RF module able to communicate through SPI communication 10](#__RefHeading___Toc1113_2042270136)

[3 Drone controller board 12](#__RefHeading___Toc1115_2042270136)

[3.1 Pinout 12](#__RefHeading___Toc1118_2042270136)

# Table of Figures

[*Figure 1: Top and bottom layer of drone main board* 5](#Figure!0|sequence)

[*Figure 2: Battery Connector* 7](#Figure!1|sequence)

[*Figure 3: Motor power connectors* 7](#Figure!2|sequence)

[*Figure 4: Motor control connectors* 8](#Figure!3|sequence)

[*Figure 5: SD card connector & module* 8](#Figure!4|sequence)

[*Figure 6: BME/BMP280 connector & module* 9](#Figure!5|sequence)

[*Figure 7: MPU6050 connector & module* 9](#Figure!6|sequence)

[*Figure 8: GPS GT-U7 connector & module* 10](#Figure!7|sequence)

[*Figure 9: RF module connectors & module* 10](#Figure!8|sequence)

[Figure 10: Ultrasonic sensor connectos & module 11](#Figure!9|sequence)

[*Figure 11: Bluetooth connector & module* 11](#Figure!10|sequence)

# Table of table

[Table 1 Expansion board pinout 5](#_Toc20988)

# 1 Introduction

## **1.1 Overview of Sound Recognition Technology**

Sound recognition technology allows electronic devices to interpret and respond to audio signals, enabling voice commands, environmental sound detection, and real-time interaction. This technology is widely used in applications such as **voice assistants (e.g., Alexa, Google Assistant), security systems, and accessibility tools** for individuals with disabilities.

At its core, sound recognition involves **capturing audio signals, processing them to extract features, and matching them to predefined patterns**. With the advancement of machine learning and edge computing, microcontrollers like the **ESP32** can now perform sound recognition efficiently without requiring cloud-based processing.

## **1.2 Importance of Embedded Sound Recognition**

Embedded sound recognition plays a crucial role in creating smart and interactive systems. Unlike cloud-based solutions, **embedded sound recognition processes audio locally on the device**, offering advantages such as:

* **Faster response time** (no internet dependency)
* **Better privacy and security** (audio data is not sent to external servers)
* **Lower power consumption** (suitable for battery-powered devices)

For students, learning **embedded sound recognition** provides valuable insights into **digital signal processing (DSP), machine learning on microcontrollers, and real-world embedded system development**.

### **1.3 Objectives of This Project**

The goal of this project is to provide **students with hands-on experience** in implementing a **sound recognition system** using the **ESP32 microcontroller**. This system will:

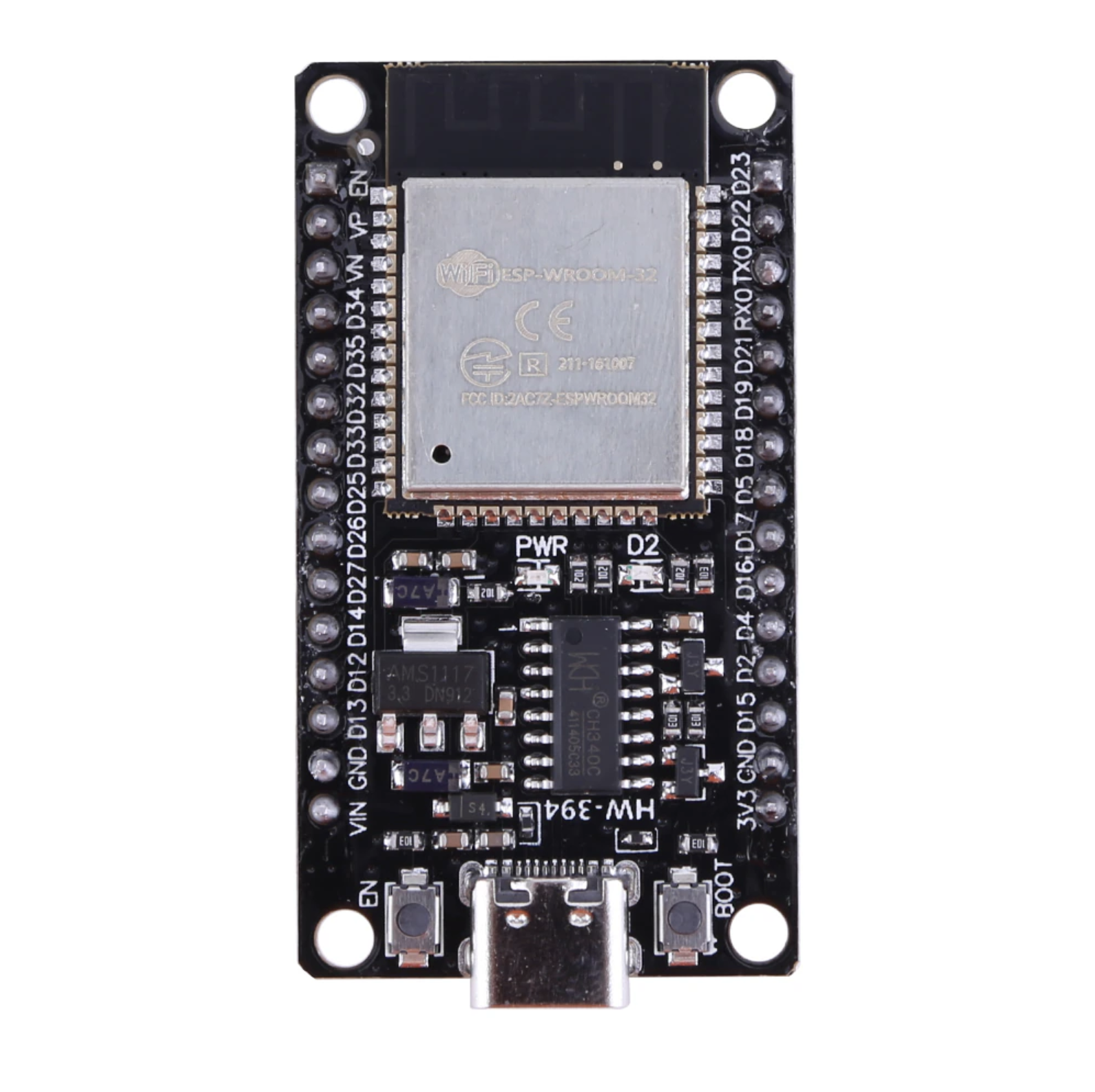
1. **Recognize voice commands** using the **VC-02 voice module**
2. **Retrieve and announce temperature & humidity** data from an **SHT31 sensor**
3. **Provide audio responses** via **I2S speaker output**
4. **Display real-time readings** on an **SSD1306 OLED screen**
5. **Use LittleFS for storing and playing pre-recorded audio files**

By completing this project, students will learn about **microcontroller programming, interfacing with hardware components, and implementing embedded sound recognition in practical applications**.

# **2. Hardware Selection**

The selection of hardware components plays a crucial role in ensuring smooth operation and performance of the embedded sound recognition system. Each component has been chosen to provide efficient processing, reliable voice recognition, clear audio output, real-time display feedback, and accurate environmental sensing.

## **2.1 ESP32 – Microcontroller for Processing**

Figure 1: ESP32 Dev Module Type C

The **ESP32** is the core processing unit of this project, responsible for handling all data processing, communication, and control operations. It is chosen due to its:

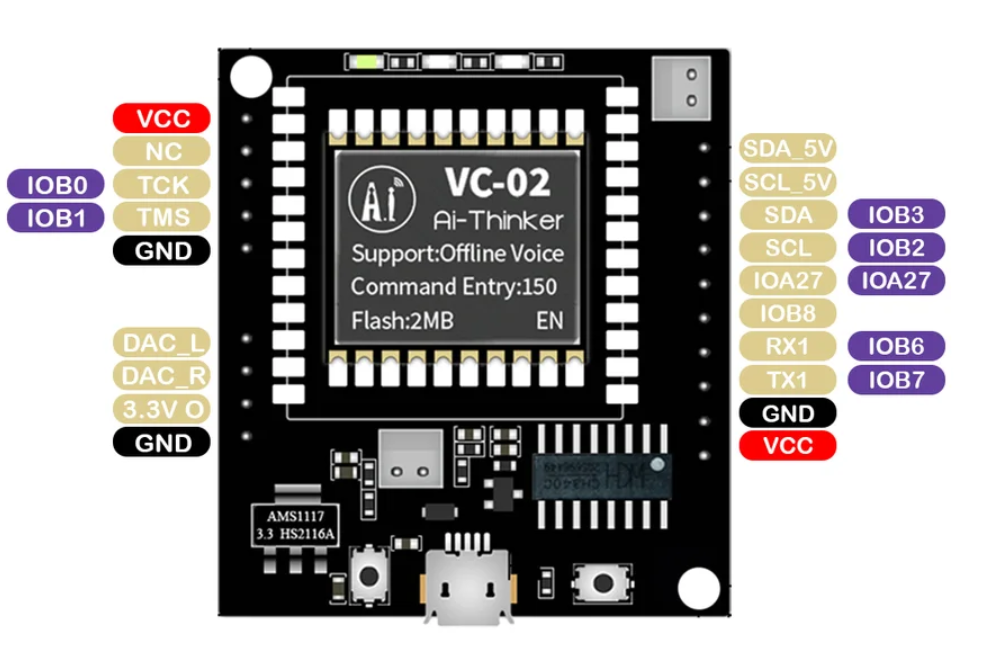
* **Dual-core processor** for handling multiple tasks efficiently
* **Wi-Fi & Bluetooth capabilities** (though not used in this project, they allow future expansion)
* **Integrated I2S, I2C, and UART interfaces** for seamless connection with peripherals
* **Low-power operation** suitable for battery-powered applications

The ESP32 manages the **voice recognition module (VC-02)**, processes **temperature and humidity sensor readings (SHT31)**, controls the **OLED display (SSD1306)**, and plays **audio responses using the MAX98357A I2S audio amplifier**.

| Component | ESP32 Pin | Function |
| --- | --- | --- |
| VC-02 TX | GPIO16 | UART RX (Receive voice commands) |
| VC-02 RX | GPIO17 | UART TX (Send data if required) |
| I2S BCLK (MAX98357A) | GPIO25 | I2S Bit Clock |
| I2S LRC (MAX98357A) | GPIO26 | I2S Left-Right Clock |
| I2S DOUT (MAX98357A) | GPIO27 | I2S Data Output |
| SSD1306 OLED SDA | GPIO21 | I2C Data |
| SSD1306 OLED SCL | GPIO22 | I2C Clock |
| SHT31 SDA | GPIO21 | I2C Data (Shared with OLED) |
| SHT31 SCL | GPIO22 | I2C Clock (Shared with OLED) |

Table 1: ESP32 Pin Connection

## **2.2 VC-02 Voice Recognition Module – Recognizing Voice Commands**

Figure 2: VC-02 Module

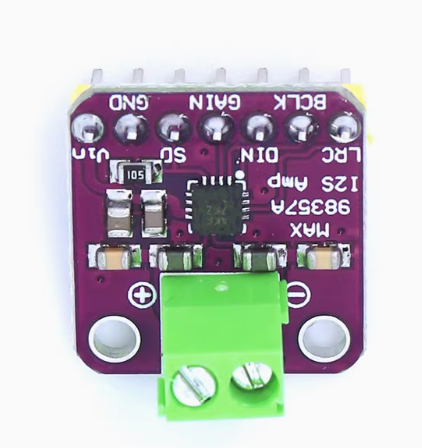
The **VC-02** module is a compact, pre-trained voice recognition module capable of recognizing custom voice commands. It is selected due to:

* **Built-in offline voice recognition** (no need for internet access)
* **Customizable wake words & commands**
* **UART interface for easy communication with ESP32**
* **Low power consumption**

### **Function in the Project**

* Detects specific voice commands like "Hi Pudding", "Humidity", “Temperature” and "Siren"
* Sends recognized command data to ESP32 via **UART**

## **2.3 I2S Audio Output (MAX98357A + Speaker) – Playing Responses**

Figure 3: MAX98357A

The **MAX98357A** is a **digital-to-analog converter (DAC) and audio amplifier** that converts digital I2S signals into analog audio output, which is played through a **connected speaker**.

### **Reasons for Choosing MAX98357A**

* **I2S support** for high-quality audio playback
* **Built-in amplifier** eliminates the need for additional circuitry
* **Compact and low-power** design

### **Function in the Project**

* Receives **pre-recorded WAV audio files** stored in ESP32's **LittleFS** storage
* Converts digital audio signals into analog signals
* Outputs sound through a connected speaker

## **2.4 SSD1306 OLED Display – Visual Feedback**

Figure 4: SSD1306 OLED Display

The **SSD1306 OLED display** is a **128x64 pixel monochrome display** used to **show real-time sensor readings and system status**.

### **Reasons for Choosing SSD1306**

* **Low power consumption** (suitable for embedded applications)
* **I2C communication** (only requires 2 pins, SDA & SCL)
* **Clear, high-contrast display** for easy readability

### **Function in the Project**

* Displays **real-time temperature and humidity readings**
* Provides **visual confirmation** of recognized voice commands

## **2.5 SHT31 Temperature & Humidity Sensor – Sensor Data Acquisition**

The **SHT31** is a highly accurate **temperature and humidity sensor** that communicates with the ESP32 over **I2C**.

### **Reasons for Choosing SHT31**

* **High accuracy** (±0.3°C for temperature, ±2% for humidity)
* **Fast response time**
* **I2C communication** (easy interface with ESP32)

### **Function in the Project**

* Measures **temperature and humidity** in real-time
* Sends data to the ESP32 for processing and display
* Enables the **audio response system** to announce current weather conditions