INTERNATIONAL JOURNAL OF RESEARCH AND SCIENTIFIC INNOVATION (IJRSI)
ISSN No. 2321-2705 | DOI: 10.51244/IJRSI |Volume XIII Issue VII July 2026

Design and Implementation of an AI Emotion Recognition-Based Smart Diffuser System for Vehicles

Yonghwan Son[1], Jaeseong Kim[2], Dohun Lee[3*], Seungjae Lee[4*]

[1,2,3,4]Department of Computer Engineering, Sun Moon University, Korea

***Corresponding Author**

DOI: https://doi.org/10.51244/IJRSI.2026.1307000266

Received: 28 July 2026; Accepted: 03 August 2026; Published: 13 August 2026

ABSTRACT

This paper proposes an AI emotion recognition-based smart diffuser system for vehicles that monitors a driver's emotional and drowsiness states in real time, dynamically controlling fragrance diffusion, visual feedback, and auditory alerts. Traditional vehicle diffusers are typically limited to manual operation or fixed periodic spraying, making it difficult to actively adapt to changing driver conditions. To overcome these limitations, the proposed system integrates a Raspberry Pi-based AI recognition unit, an ESP32-based hardware control unit, a Flask REST API server, and a React Native mobile application. The Raspberry Pi captures camera frames to detect the driver’s face and analyze emotions and drowsiness, while the ESP32 manages RGB LEDs, fragrance diffusion modules, a cooling fan, a buzzer, and an OLED display via UART commands. Experimental results demonstrate that the emotion classification model achieves an accuracy of 83.46%, and the drowsiness detection module achieves over 90% accuracy. These findings validate that the proposed system serves as an effective, intelligent in-vehicle environment control device capable of actively responding to the driver's physiological and emotional states.

Keywords: emotion recognition, smart diffuser, ESP32, drowsiness detection, in-vehicle environment control

INTRODUCTION

A driver's emotional and drowsiness states are critical factors directly linked to driving safety. Conditions such as anger, anxiety, sadness, fatigue, and drowsiness significantly diminish concentration and reaction times, with drowsy driving remaining a leading cause of severe traffic accidents[1],[2]. Consequently, developing technologies that continuously assess a driver's condition in real time and actively regulate the in-vehicle environment has become increasingly essential[3],[4].

Advancements in artificial intelligence and embedded systems now enable compact computing devices, such as the Raspberry Pi, to efficiently execute image-based AI tasks including face detection, facial expression classification, and eye-state analysis[5]. Furthermore, IoT microcontrollers like the ESP32 provide robust capabilities for controlling diverse sensors and actuators, bridging computational intelligence with physical feedback mechanisms[6]. While conventional vehicle diffusers are widely used to enhance cabin comfort, most existing products depend entirely on direct user intervention or simple, repetitive spraying cycles[7]. These passive approaches fail to reflect real-time shifts in a driver's emotional or fatigue levels, rendering them inadequate for proactive intervention[8]. To address these shortcomings, this study introduces an intelligent in-vehicle smart diffuser system that combines AI-based emotion recognition with IoT hardware control[9]. By capturing the driver's facial expressions via camera, the system estimates emotional states and drowsiness levels to automatically regulate fragrance modules, RGB LEDs, cooling fans, buzzers, and OLED displays.

Ultimately, this research aims to transcend the boundaries of conventional passive diffusers by delivering an adaptive, safety-oriented in-vehicle environment control solution[10].

PROPOSED SYSTEM

System Architecture

The proposed smart diffuser adopts a three-layer architecture consisting of a Raspberry Pi 4, an ESP32 LoLin D32 Pro, and a user smartphone application named AromaCare. The overall system architecture is shown in Figure 1.

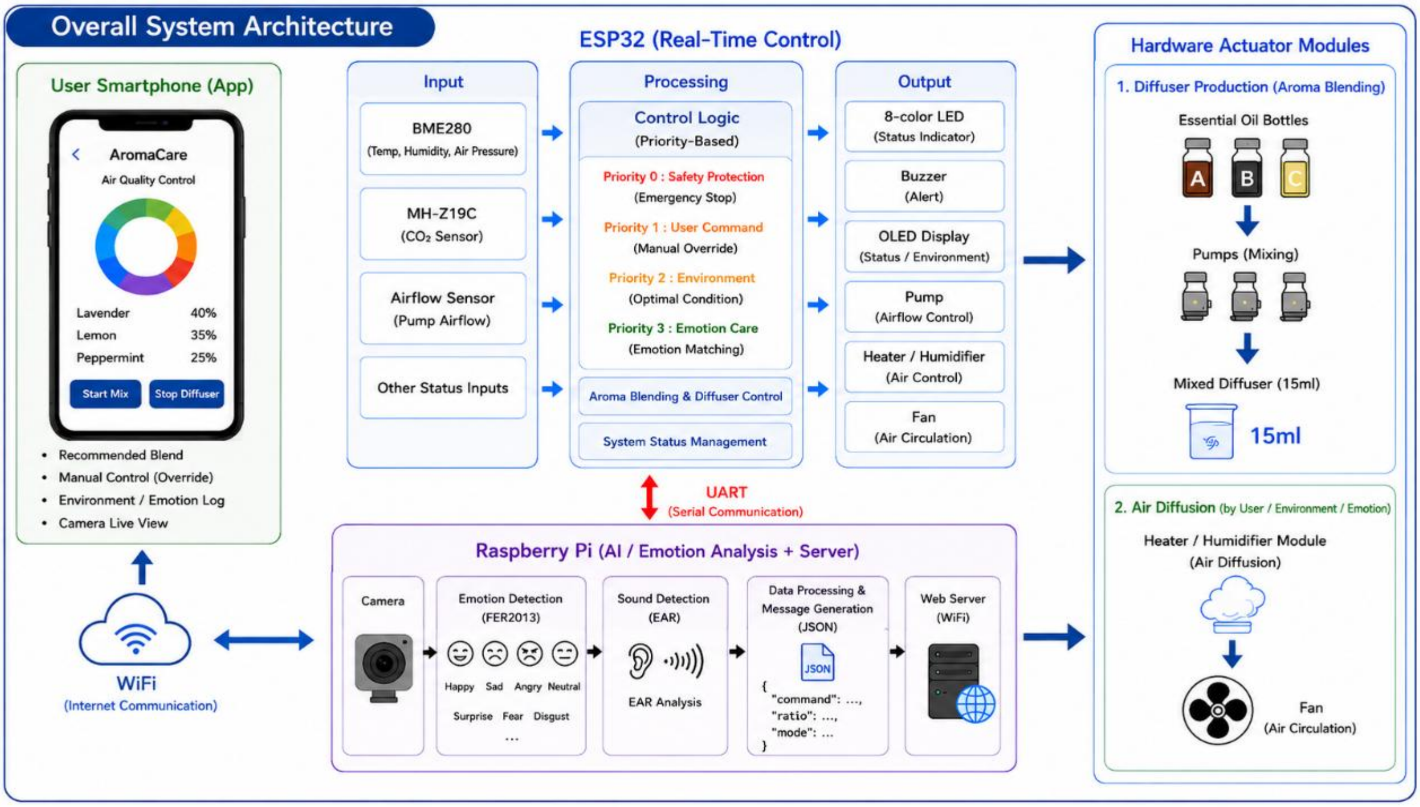

Figure 1. Overall system architecture.

Facial expression images collected through the camera are analyzed on the Raspberry Pi 4 using an emotion recognition model based on FER2013. The analyzed emotion result is converted into a structured command and transmitted to the ESP32 through bidirectional UART communication. The ESP32 integrates environmental data from BME280 and MH-Z19C sensors and controls the PWM intensity of the top, middle, and base-note ultrasonic diffuser modules. The mobile application connects to the Raspberry Pi server through Wi-Fi and provides fragrance recommendation, manual control, environmental and emotion logs, and real-time camera monitoring.

Priority-Based Control Logic

The ESP32 firmware operates according to a four-level priority control policy. This policy ensures that safety-related actions are processed before comfort-oriented or emotion-care actions.

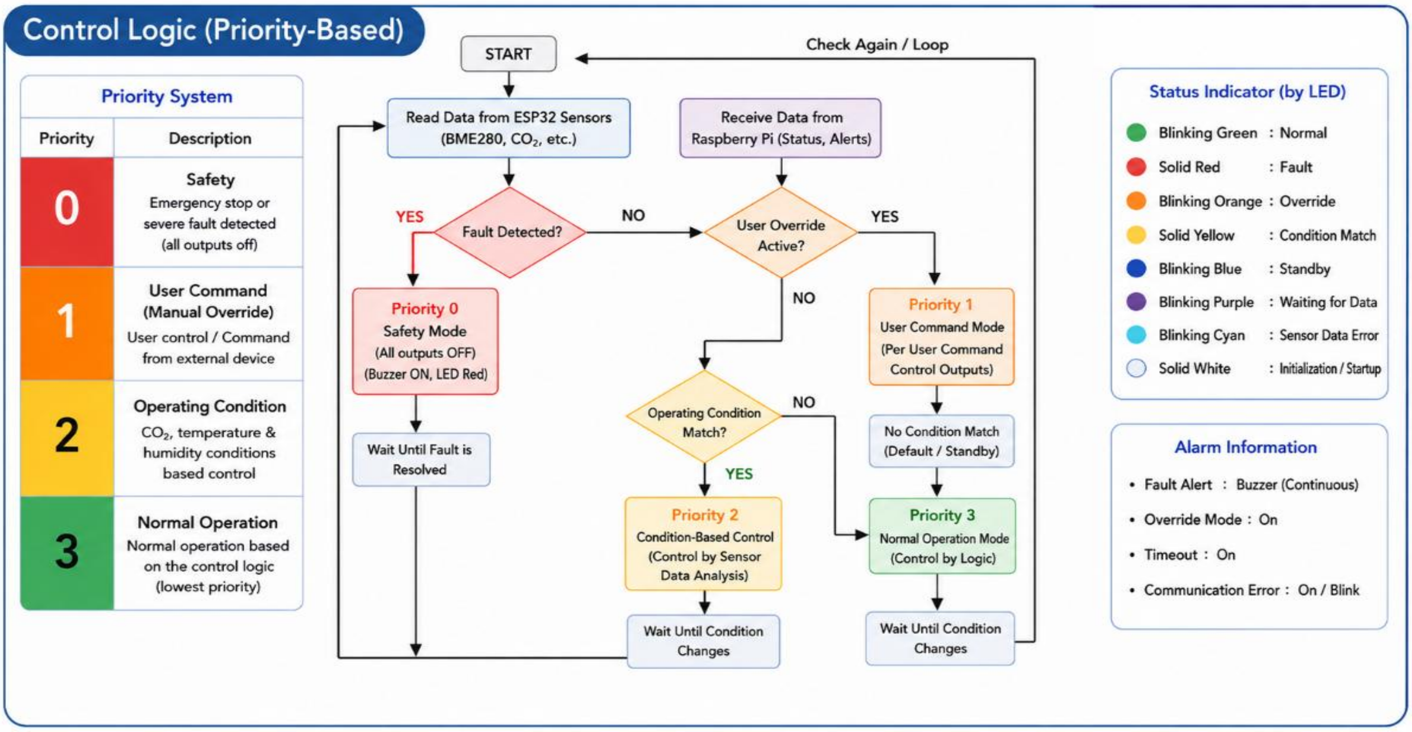

Figure 2. Priority-based control logic.

Priority 0, life protection, has the highest priority. When drowsiness is detected, the system immediately interrupts all lower-level operations and generates an alert using the buzzer and red LED. The drowsiness state is determined by eye-state analysis using the Eye Aspect Ratio (EAR).

Priority 1 processes user manual commands from the smartphone application. Manual fragrance ratio settings, diffusion start and stop commands, and direct channel control are processed before automatic emotion-care operation.

Priority 2 optimizes the vehicle interior environment by using CO2 concentration, temperature, and humidity data measured by the BME280 and MH-Z19C sensors. When CO2 concentration exceeds 1,500 ppm, the top-note output is increased, and PWM compensation is applied according to temperature and humidity thresholds.

Priority 3 is the lowest-priority emotion-care mode. In this mode, the system performs fragrance diffusion according to the recognized emotional state. The PWM ratios for top, middle, and base notes are defined in Table 1.

Table 1. Emotion-based fragrance diffusion mapping.

Emotion value	Emotion state	Top note (GPIO 32)	Middle note (GPIO 33)	Base note (GPIO 25)	Mood
3	Happy	220	150	80	Bright and cheerful
0	Angry	100	220	150	Calming and relaxing
4	Sad	80	220	150	Comforting and stable
6	Neutral	150	150	150	Balanced and neutral
-1	Idle	0	0	0	Standby

Fragrance Blending and Diffusion System

The fragrance materials used in this system are divided into three groups according to perfume composition principles: top note, middle note, and base note. Each group is prepared as a pre-blended solution and mixed with a diffuser base so that it can be applied to the ultrasonic humidifier module with low viscosity.

Table 2. Fragrance note composition.

Note	Fragrance composition	Characteristics
Top note	Lemon, orange, bergamot	Initial scent, fast evaporation, citrus family
Middle note	Lavender, chamomile, jasmine	Core scent, medium evaporation, floral family
Base note	Cedarwood, sandalwood, white musk	Residual scent, slow evaporation, woody family

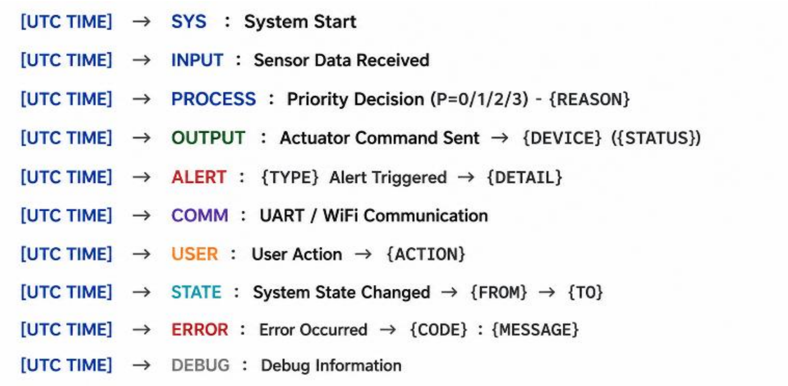

Figure 3. Diffusion and blending sequence.

Three ultrasonic humidifier modules are filled with the top, middle, and base-note solutions. The generated mist is combined inside an acrylic mixing chamber so that the fragrance can be blended through collision and airflow mixing. The cooling fan assists diffusion toward the user.

Table 3. Environmental compensation logic.

Environmental condition	Compensation	Purpose
CO2 > 1,500 ppm	Top note +30	Increase refreshing impression
Temperature > 26°C	Top note +20	Increase cool impression
Temperature < 18°C	Base note +20	Increase warm impression
Humidity < 40%	All channels +20	Increase diffusion intensity
Humidity > 70%	All channels -20	Reduce diffusion intensity

Emotion Recognition Model

A transfer-learning model based on MobileNetV2 is used for emotion recognition. The model receives a 160 × 160 RGB face image and produces a four-class softmax output. The Raspberry Pi 4 captures camera frames in real time, performs face detection, and simultaneously executes EAR-based drowsiness detection. The recognition result is converted into a command packet and transmitted to the ESP32 at five-second intervals.

Hardware Configuration

The main hardware components used in the proposed system are summarized in Table 4. An IRLZ44N MOSFET switching circuit allows the ESP32 3.3 V GPIO signals to control the diffuser modules and cooling fan through PWM. Each channel includes a 1N5819 flyback diode and a 10 kΩ gate-source pulldown resistor to protect the GPIO pins from reverse voltage and unintended switching.

Table 4. Main hardware components.

Category	Component	Model	Purpose
MCU	ESP32 board	LoLin D32 Pro	Dedicated hardware control
MCU	Single-board computer	Raspberry Pi 4	Emotion recognition and server
Sensor	Temperature/humidity/pressure sensor	BME280	Environmental monitoring (I2C)

ISSN No. 2321-2705 | DOI: 10.51244/IJRSI |Volume XIII Issue VII July 2026

Sensor	CO2 sensor	MH-Z19C	CO2 concentration measurement (UART)
Output	Ultrasonic humidifier module	SZH-HUM003	Fragrance diffusion (PWM control)
Output	RGB LED stick	WS2812B (8 LEDs)	Emotion-state visualization
Output	OLED display	SSD1306 0.96 inch	Real-time status display
Output	Piezo buzzer	SFM-27	Alert sound
Output	Cooling fan	5 V, 50 mm	Fragrance diffusion
Control	MOSFET	IRLZ44N	PWM switching, 3.3 V logic compatible
Protection	Diode	1N5819	Reverse-voltage protection

Software Flow

The software is organized into four modules. First, the ESP32 collects environmental sensor data from the BME280 and MH-Z19C at five-second intervals, displays the current status on the OLED, and transmits the data to the Raspberry Pi through UART. Second, the Raspberry Pi performs camera acquisition, face detection, emotion recognition, EAR-based drowsiness detection, status code generation, UART command transmission, and Flask server operation. Third, the smartphone application connects to the Raspberry Pi through Wi-Fi and provides real-time camera streaming, environmental logs, emotion and drowsiness display, fragrance recommendation, ratio setting, and manual command transmission. Finally, the ESP32 parses the received UART commands, executes priority-based control, controls fragrance diffusion, updates OLED and LED states, and stores or retransmits logs when necessary.

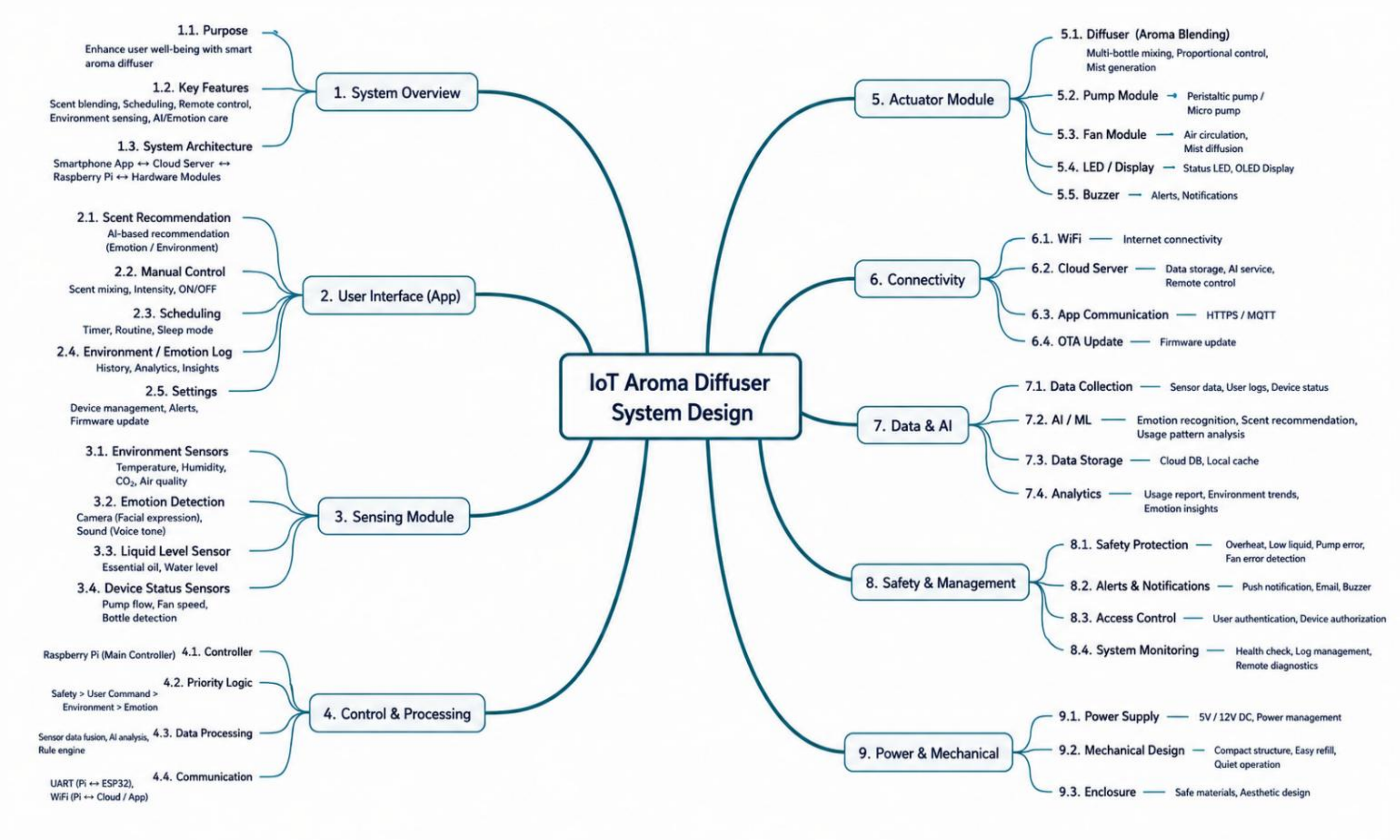

Figure 4. IoT Aroma Diffuser System Design.

Experiments and Evaluation

To evaluate the proposed AI-based smart diffuser system, experiments were conducted for emotion recognition, drowsiness detection, environmental sensor accuracy, PWM-based diffusion control, and Raspberry Pi-ESP32 integration.

Table 5. Training data configuration by emotion class.

Emotion class	Train	Validation
Angry	8,390	1,390
Happy	8,000	1,818
Sad	8,000	1,517
Neutral	8,000	1,530

Emotion Recognition and Drowsiness Detection

The emotion classification model was trained to classify four emotion classes: Happy, Angry, Sad, and Neutral. After face detection, the face region was preprocessed to match the model input size. The training dataset was configured by limiting the number of samples per class in order to reduce class imbalance.

Table 6. Model performance.

Model	Input data	Classes	Accuracy
Emotion classification model	Face image	4	83.46%
Drowsiness detection module	Eye-region image / eye-state information	2	More than 90%

Table 7. Average recognition accuracy by emotion.

Emotion	Average accuracy
Happy	88%
Neutral	85%
Sad	80%
Angry	78%

The overall average accuracy measured in the smart diffuser environment was approximately 82.75%. Stable recognition performance was observed under sufficient lighting and frontal-face conditions. However, partial accuracy degradation occurred when the face was turned sideways or the lighting was insufficient.

Environmental Sensor Accuracy

Table 8. Environmental sensor measurement accuracy.

Measurement item	Reference value	Measured value	Error
Temperature	23.0°C	23.2°C	±0.5°C
Humidity	54%	55%	±2%
Pressure	1013 hPa	1012.8 hPa	±1 hPa

The BME280 sensor measurements were within the acceptable error range compared with a commercial digital thermo-hygrometer. The MH-Z19C CO2 sensor provided stable measurements after warm-up. In a general indoor environment, the measured concentration was approximately 400-600 ppm, while in an enclosed environment it reached approximately 1,200-1,500 ppm. When the CO2 concentration exceeded 1,500 ppm, the environmental compensation algorithm increased the top-note PWM value by 30 as intended.

PWM-Based Diffusion Control

Table 9. PWM control performance.

PWM duty	Diffusion state
0	Stopped
64	Very weak
128	Weak

192	Medium
220	Strong
255	Maximum

Table 10. Diffusion sequence timing.

Step	Configured time	Measured time
Diffusion ON	5000 ms	5012 ms
Transition delay	500 ms	510 ms
Fan ON	5000 ms	5008 ms
Total cycle	10500 ms	10530 ms

The ultrasonic diffuser did not generate stable mist below the effective duty threshold. Therefore, the minimum stable PWM duty value was set to 220 in the implemented scaling algorithm. The total timing error of the diffusion sequence was within ±30 ms, satisfying the design target.

Integrated Control Verification

UART communication tests were performed to confirm whether AI recognition results from the Raspberry Pi were correctly delivered to the ESP32 hardware control unit. The ESP32 changed the LED color, operated the fragrance diffusion modules, controlled the fan, and generated buzzer alerts according to the received command packets.

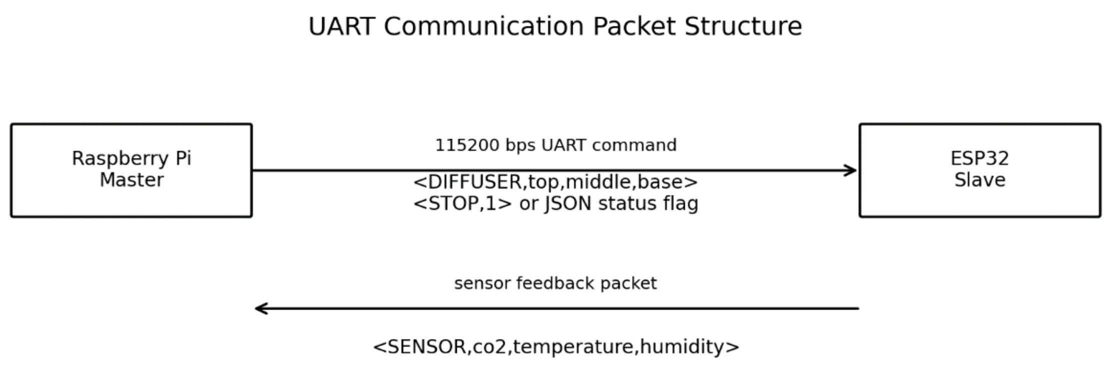

Figure 5. UART communication packet structure.

Table 11. Priority-control verification results.

Scenario	Result
Drowsiness detected	Emergency mode entered correctly
Manual app control	Automatic control interrupted correctly
CO2 environmental compensation	Top-note compensation performed correctly
Emotion-based diffusion	Fragrance combination executed correctly
Priority conflict	Highest-priority action performed correctly

LED, OLED, and Mobile Application Verification

The WS2812B LED and SSD1306 OLED were tested as user-feedback interfaces. The LED displayed different colors according to emotion state, while the OLED showed emotion state, CO_2 concentration, temperature, humidity, current blend, and PWM intensity in real time. The mobile application communicated with the Raspberry Pi server through Wi-Fi-based HTTP requests and provided control functions such as diffusion intensity adjustment, automatic/manual mode switching, and system status inquiry.

ISSN No. 2321-2705 | DOI: 10.51244/IJRSI |Volume XIII Issue VII July 2026

Table 12. Mobile application control verification.

Function	Result
Automatic mode activation	Normal
Manual mode activation	Normal
Diffusion start	Normal
Diffusion stop	Normal
Top-note intensity change	Normal
Middle-note intensity change	Normal
Base-note intensity change	Normal
System status inquiry	Normal

Table 13. Overall performance evaluation.

Evaluation item	Result
Emotion recognition accuracy	82.75%
Drowsiness detection accuracy	More than 90%
Environmental sensor error	Within ±2%
Diffusion timing error	Within ±30 ms
UART communication stability	Normal
Priority-control success rate	100%
LED/OLED status display	Normal

IMPLEMENTATION

Hardware Module Fabrication and Power Distribution

The proposed system is a complex embedded system that combines an AI inference module requiring a relatively large amount of computation and an actuator control module that operates sensors and output devices in real time. To secure reliability in the limited space and power environment of a vehicle, the hardware was implemented using functional module separation and noise-aware power design.

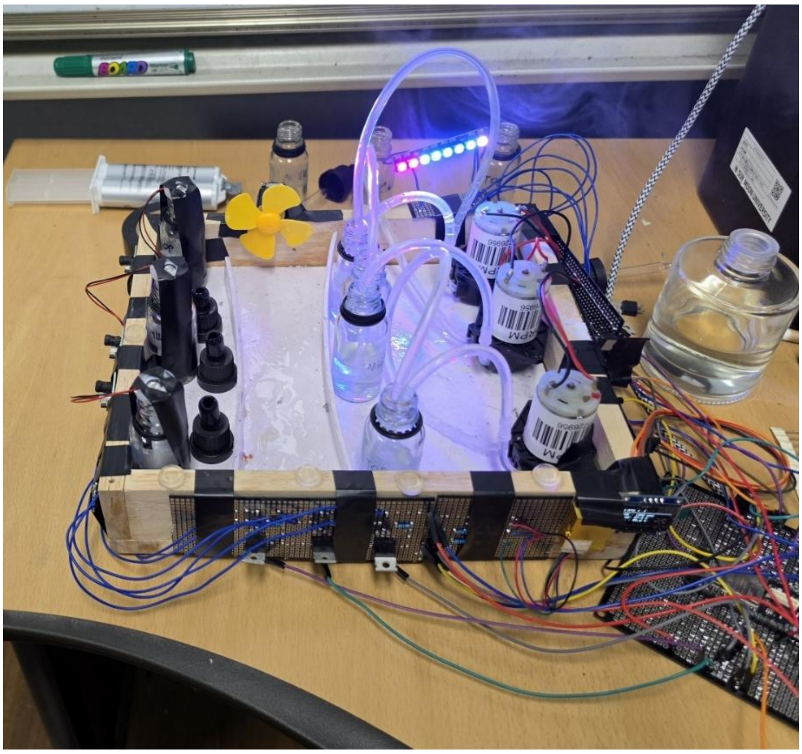

Figure 6. Fabricated smart diffuser prototype.

Figure 7. Separated power distribution and common ground design.

The hardware was divided into a main control and power board and a sensor/display interface board. The main board integrates the Raspberry Pi 4, ESP32 MCU, MOSFET switching circuits, and power lines for high-current components. The sensor/display board places the OLED display, RGB LED, and environmental sensors on a separate slim board to improve installation convenience. High-current devices are powered directly from the 5 V output of the DC-DC buck converter, while noise-sensitive components use the regulated 3.3 V logic supply. All modules share a common ground to prevent UART signal reference distortion.

UART Serial Communication and Firmware Logic

Data exchange between the Raspberry Pi 4 master and ESP32 slave is performed using UART serial communication at 115200 bps. Instead of using an unstructured byte stream, the system uses JSON or structured packet formats to improve extensibility and command clarity. When the Raspberry Pi sends a packet containing emotion codes and status flags, the ESP32 parses the packet and controls the lower-level actuators.

ISSN No. 2321-2705 | DOI: 10.51244/IJRSI |Volume XIII Issue VII July 2026

Table 14. JSON-based UART communication between master and slave.

Direction	Communication protocol / packet	Purpose
Raspberry Pi → ESP32	UART2, 115200 bps: {"cmd":"aroma"}	Automatic diffusion command
Raspberry Pi → ESP32	UART2, 115200 bps: {"cmd":"manual","top":15,"mid":40,"base":20}	Manual app command
Raspberry Pi → ESP32	UART2, 115200 bps: {"cmd":"emotion","value":2}	Emotion-state command
ESP32 → Raspberry Pi	UART2, 115200 bps: {"co2":1350,"temp":24.5,"status":"diffusing"}	Sensor and status feedback

To improve fragrance-blending accuracy, the firmware compensates for the physical nonlinearity and dead zone of the DC pump and ultrasonic diffuser. For the ultrasonic diffuser, logical diffusion intensity from 1% to 100% is mapped into the effective operating PWM range. For liquid blending, the pump motor voltage is fixed at maximum output and the operating time is controlled using millis() instead of delay(). This non-blocking control allows the three pumps for top, middle, and base notes to stop independently when each target blending time is reached.

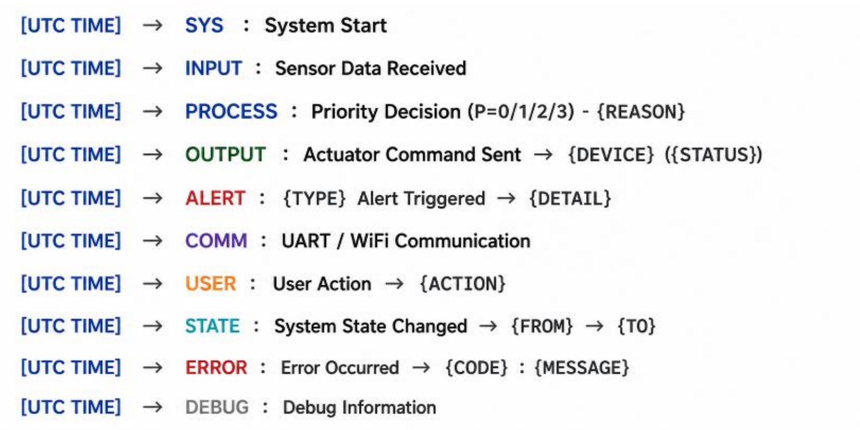

Figure 8. Pump control log and PWM conversion monitoring result.

Mobile Application and User Interface

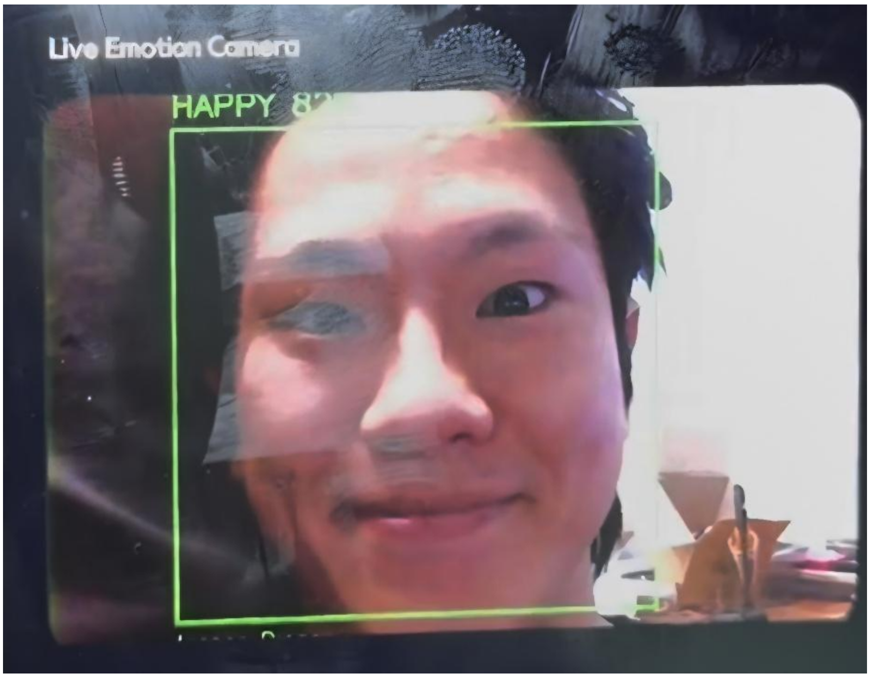

Figure 9. Real-time emotion recognition screen.

ISSN No. 2321-2705 | DOI: 10.51244/IJRSI |Volume XIII Issue VII July 2026

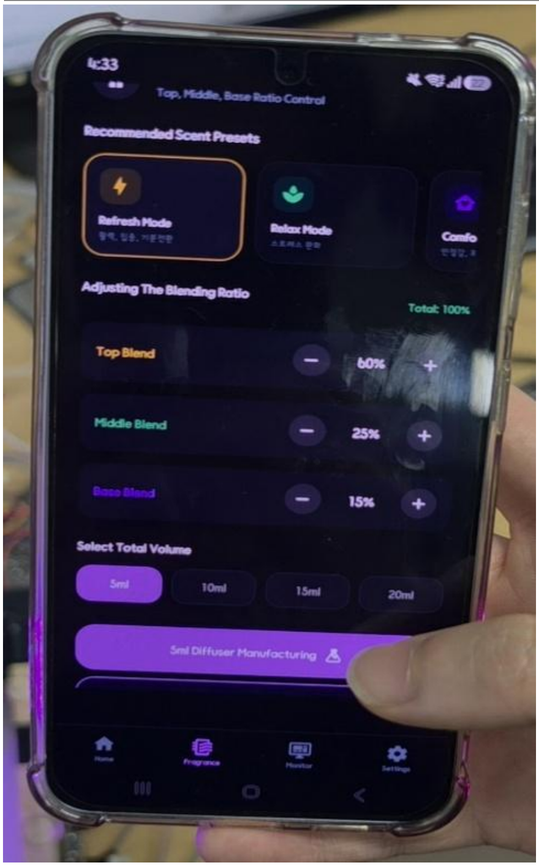

Figure 10. Mobile application control screen.

The mobile application allows the user to check the current system status and switch between automatic and manual modes. In automatic mode, the diffuser operates based on AI recognition results. In manual mode, the user can directly select top, middle, and base channel values. The Flask REST API server runs on the Raspberry Pi and handles communication between the mobile application and the Raspberry Pi. The Raspberry Pi converts received HTTP requests into UART packets and forwards them to the ESP32. Sensor data can also be transmitted from the ESP32 to the Raspberry Pi using the <SENSOR,co2,temperature,humidity> packet format so that vehicle interior environmental data can be displayed in the application.

Table 15. Implementation environment of the proposed system.

Item	Description
AI inference / image processing	Raspberry Pi, Python, OpenCV, TensorFlow Lite
Hardware control	ESP32 LoLin D32 Pro, Arduino C++
Emotion classification	Face image-based emotion classification model
Drowsiness detection	Eye-state analysis and sustained eye-closure detection
Communication	UART, Flask REST API, HTTP
Mobile application	React Native
Controlled devices	RGB LED, fragrance diffusion module, cooling fan, buzzer, OLED
Sensors	CO2, temperature, and humidity sensors
Model training environment	Kaggle GPU environment

CONCLUSION

This paper designed and implemented an AI emotion recognition-based smart diffuser system for vehicles using Raspberry Pi This paper successfully designs and implements an AI emotion recognition-based smart diffuser system for vehicles utilizing a Raspberry Pi and an ESP32. The system tracks the driver's real-time emotional and drowsiness status, dynamically coordinating fragrance diffusion, visual cues, and auditory warnings accordingly.

By integrating camera-based emotion recognition, eye-state-based drowsiness detection, UART-driven hardware management, and a Flask REST API mobile application, the system delivers comprehensive in-vehicle interaction. Experimental evaluations confirmed an 83.46% accuracy rate for the emotion classification model and over 90% accuracy for drowsiness detection. Moreover, the ESP32 control unit reliably executed emotion-tailored LED indicators, multi-note fragrance blending, airflow regulation, and emergency buzzer alerts.

Differentiating itself from traditional manual or periodic diffusers, the proposed architecture prioritizes driver safety by treating drowsiness detection as a top-priority intervention, highlighting its potential to evolve from a mere convenience amenity into a vital driver-safety assistance tool.

Because this prototype was evaluated under controlled experimental setups, future work will focus on addressing real-world driving variables—such as fluctuating ambient lighting, changing camera angles, partial facial occlusion, and vehicle vibrations. Comprehensive user studies and long-term driving evaluations will also be conducted to quantitatively assess the psychological and physiological impacts of customized fragrance diffusion on drivers. Subsequent research directions include enhancing face recognition robustness, refining emotion classification accuracy, integrating vehicle OBD data, supporting multi-user recognition, optimizing scent-blending algorithms, and deploying extended tests in actual automotive environments.

ACKNOWLEDGMENT

This research was supported by the MSIT(Ministry of Science ICT), Korea, under the National Program for Excellence in SW, supervised by the IITP(Institute of Information & Communications Technology Planning & Evaluation) in 2026 (No. 2024-0-00023).

REFERENCES

1. M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L. Chen, "MobileNetV2: Inverted Residuals and Linear Bottlenecks," Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 4510-4520, 2018.
2. S. Li and W. Deng, "Reliable Crowdsourcing and Deep Locality-Preserving Learning for Facial Expression Recognition," IEEE Transactions on Image Processing, vol. 28, no. 1, pp. 356-370, 2019.
3. A. Mollahosseini, B. Hasani, and M. H. Mahoor, "AffectNet: A Database for Facial Expression, Valence, and Arousal Computing in the Wild," IEEE Transactions on Affective Computing, vol. 10, no. 1, pp. 18-31, 2019.
4. P. Viola and M. Jones, "Rapid Object Detection using a Boosted Cascade of Simple Features," Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 511-518, 2001.
5. TensorFlow, "TensorFlow Lite Guide," Google Developers.
6. OpenCV, "Open Source Computer Vision Library Documentation."
7. Espressif Systems, "ESP32 Technical Reference Manual."
8. Flask Documentation, "Flask Web Development Framework."
9. React Native Documentation, "Learn Once, Write Anywhere."
10. Raspberry Pi Foundation, "Raspberry Pi Documentation."