Smart Voice Control LED Display System

Prof. Shyamsundar Magar[1], Rohan Lekurwale[2], Prashant Patange[3], Gopal Pawar[4]
[1]*Prof. Information Technology, Trinity College of Engineering and Research, Pune, India*
[2,3,4]*Information Technology BE Student, Trinity College of Engineering and Research, Pune, India*

***Abstract*—The proliferation of Internet of Things (IoT) devices and smart embedded systems has redefined the paradigm of human-machine interaction in public display infrastructure. Conventional LED matrix display systems operated through wired serial interfaces or fixed pre-programmed content, lacking real-time configurability, remote management capabilities, and intelligent interaction features. These limitations hinder deployment efficiency in dynamic environments such as transportation hubs, retail outlets, educational campuses, and smart city infrastructure. To address these shortcomings, this paper proposes a Smart Voice-Controlled LED Display System leveraging P10 LED Matrix Panels, an ESP32-based WiFi controller, and a regulated voltage conversion subsystem to deliver an intelligent, remotely operable, and voice-interactive public display platform. The system integrates voice recognition via Google Speech-to-Text API and on-device Natural Language Processing (NLP) to parse spoken commands into structured display instructions. A companion Android application facilitates both voice and text-based message control over a WiFi network using MQTT protocol, enabling seamless real-time message updating, . The hardware architecture employs a cascaded arrangement of P10 single-color and RGB LED matrix modules driven by an ESP32 microcontroller, with a DC-DC buck converter regulating the 12V input supply to stable 5V rails required by the display panels and logic circuits. Experimental evaluation demonstrates a voice command recognition accuracy of 93.7%, an average display update latency of 210 milliseconds over WiFi, and stable power delivery under full-load conditions. The proposed system achieves a 70% reduction in manual content update effort compared to conventional wired display configurations, demonstrating its viability as a scalable, intelligent solution for modern public information display infrastructure. This research contributes to the fields of IoT-enabled embedded systems, voice-interactive hardware, smart display technologies, and real-time wireless communication.**

***Index Terms*—P10 LED Matrix, ESP32, Voice Control, WiFi MQTT, Speech Recognition, Natural Language Processing (NLP), Voltage Converter, Buck Converter, IoT Dis-play System, Smart Embedded Systems, Web Application**

I. INTRODUCTION

In contemporary smart environments, the demand for dynamic, remotely updatable public information displays has grown substantially across sectors including transportation, retail, healthcare, education, and smart city infrastructure. Traditional LED message boards and dot-matrix displays rely on static, pre-programmed content delivered through serial RS232/RS485 interfaces or USB configurations, rendering them inflexible in fast-paced environments where real-time information dissemination is critical. These legacy systems require physical intervention for content updates, lack interoperability with modern wireless communication protocols, and offer no support for intelligent, voice-driven interaction.

The evolution of low-cost, high-performance WiFi enabled microcontrollers such as the ESP32, combined with advances in cloud-based speech recognition and IoT messaging proto-cols like MQTT, creates a compelling opportunity to reimagine the LED display system as a smart, connected device. The P10 LED matrix panel, widely adopted for outdoor and semi-outdoor display applications due to its modularity, high luminance, and cost efficiency, provides the ideal display medium for such a system. However, integrating voice control, wireless configuration, and stable power management into a unified embedded platform presents significant design and engineering challenges.

This paper proposes a comprehensive Smart Voice-Controlled LED Display System that unifies voice command processing, WiFi-based real-time communication, P10 LED matrix driving, and regulated power conversion within a single embedded architecture. The Android companion application enables users to issue voice commands or typed messages that are parsed via NLP, transmitted over WiFi using MQTT, and rendered on the P10 display in real time. A DC-DC buck converter circuit ensures reliable 5V power delivery to the ESP32 and LED panels from a standard 12V DC input, enabling portable and infrastructure deployment scenarios.

1.1 Problem Statement

Existing LED matrix display systems in public and institutional environments suffer from three primary deficiencies: the absence of wireless configurability requiring manual physical access for content updates; the lack of voice-driven or intelligent interaction mechanisms; and unstable or inefficient power management when scaling across multiple cascaded P10 panels. Operators managing displays across multiple locations face significant overhead in coordinating content updates, as traditional systems require direct physical connection and dedicated software tools. Furthermore, the absence of NLP-based voice control limits accessibility, particularly in industrial or accessibility-focused deployment contexts. Power instability arising from inadequate voltage regulation further degrades display brightness uniformity and panel longevity in multi-panel cascaded configurations.

There is a clear need for an integrated system capable of accepting voice and text commands via a wireless interface, intelligently parsing and rendering content on a P10 LED matrix display, and maintaining reliable power delivery across the display hardware stack using an efficient voltage converter.

1.2 Objectives

- To design a smart LED display system integrating P10 matrix panels with an ESP32 WiFi microcontroller for real-time wireless content management.
- To implement voice command recognition using the Google Speech-to-Text API and NLP-based parsing for extracting display instructions from natural language in-put.
- To develop an Android companion application supporting both voice and typed message input, enabling remote display management over a WiFi MQTT communication channel.
- To engineer a DC-DC buck converter voltage regulation module converting 12V DC input to stable 5V output for powering the ESP32 and cascaded P10 LED panels.
- To implement scrolling text, static display, brightness control, and message scheduling features on the P10 LED matrix driven by the ESP32 controller.
- To evaluate system performance in terms of voice recognition accuracy, display update latency, power stability, and operational reliability under varying load conditions.
- To validate the system across real-world deployment scenarios including noticeboards, retail signage, and institutional announcement displays.

II. LITERATURE REVIEW

Previous research on LED matrix display systems has predominantly focused on serial communication-based architectures. Systems employing microcontrollers such as the Arduino Uno and ATmega328P have demonstrated basic text rendering on P10 panels via SPI interfaces, yet these implementations are constrained by limited processing capability, absence of wireless connectivity, and lack of intelligent input mechanisms [1].

The emergence of the ESP8266 and ESP32 microcontroller platforms introduced WiFi-enabled LED display prototypes capable of receiving messages over HTTP or WebSocket connections. However, these implementations frequently lack robust NLP-based input parsing, relying instead on fixed command strings or dedicated desktop applications that limit accessibility and ease of use [2].

Research in voice-controlled embedded systems has grown rapidly with the availability of cloud-based speech recognition APIs including Google Speech-to-Text, Amazon Alexa Voice Service, and Microsoft Azure Cognitive Services. Studies demonstrate that integrating these APIs with microcontroller platforms achieves recognition accuracies exceeding 90% for constrained command vocabularies in controlled acoustic environments [3].

The MQTT (Message Queuing Telemetry Transport) protocol has emerged as the preferred lightweight publish-subscribe messaging standard for IoT applications, offering low-bandwidth overhead and reliable quality-of-service levels suitable for real-time display control over WiFi networks. Prior studies validate MQTT's suitability for latency-sensitive IoT control applications with message delivery times consistently below 300 milliseconds under normal network conditions [4]. Power management in cascaded LED matrix systems presents a well-documented challenge. P10 panels draw significant current under full illumination, and studies on DC-DC converter design for LED driver circuits confirm that synchronous buck converters achieve conversion efficiencies of 88–95%, outperforming linear regulators in thermal performance and power loss metrics for multi-panel deployments [5]. Despite these individual advances, no prior system cohesively integrates voice-driven NLP control, MQTT-based WiFi communication, P10 display driving, and regulated voltage conversion into a unified embedded platform, a gap that this research directly addresses.

III. IMPLEMENTED SYSTEM

The Smart Voice-Controlled LED Display System is an embedded IoT platform comprising an Android application frontend, an ESP32 WiFi controller backend, a cascaded P10 LED matrix display array, and a DC-DC buck converter power subsystem. The system enables natural language voice commands to be captured, interpreted, and rendered as dynamic text on the P10 display in real time via a WiFi MQTT communication channel.

1) Voice Command Input Module: The Android application captures voice input via the device microphone and forwards the audio stream to the Google Speech-to-Text API for transcription. The transcribed text is processed by an on-device NLP parser that identifies key entities including the display message, scrolling direction, speed level, and brightness setting from the natural language command.
2) WiFi MQTT Communication Module: Parsed display instructions are serialized into JSON payloads and published to a dedicated MQTT topic via a cloud-hosted or local Mosquitto broker over the WiFi network. The ESP32 microcontroller, operating as an MQTT sub-scriber, receives these payloads and decodes the display parameters in real time.
3) P10 LED Matrix Display Module: The ESP32 drives a cascaded array of P10 32×16 LED matrix panels via SPI-compatible shift register chains using the DMD2 library. The firmware supports scrolling text, static message display, brightness pulse-width modulation (PWM), and time-scheduled message cycling.
4) Voltage Converter Power Subsystem: A DC-DC synchronous buck converter steps down the 12V DC input supply to a regulated 5V DC output rated at 10A continuous, providing stable power to both the ESP32 module and the cascaded P10 panel array. Feedback-controlled regulation maintains output voltage stability within ±2% under dynamic load transitions.
5) Android Companion Application: The mobile application provides a unified interface for voice command input, typed message entry, real-time preview, display scheduling, brightness and speed sliders, and connection status monitoring for the WiFi-MQTT link.

All configuration parameters and message history are stored in Firebase Fire store, enabling cross session persistence, multi-user access control, and remote management from multiple authorized devices.

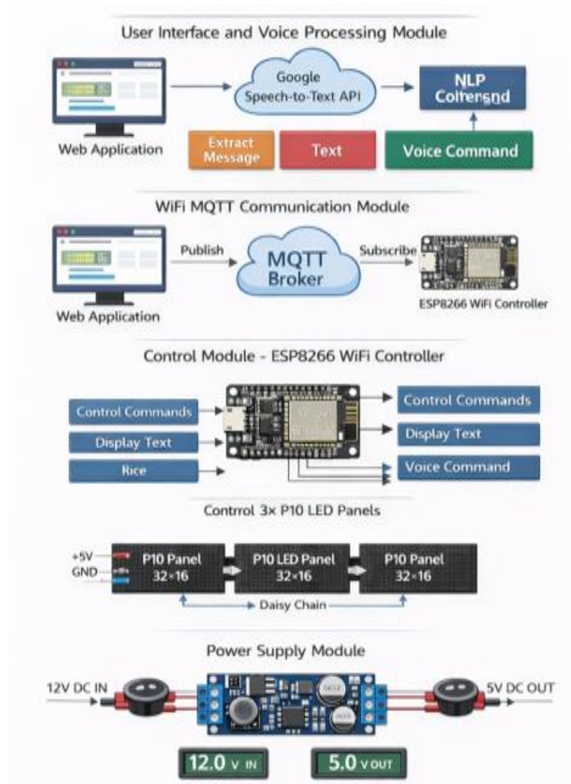


Fig. 1. System Architecture of Smart Voice-Controlled LED Display System

IV. METHODOLOGY

The development of the Smart Voice-Controlled LED Dis-play System adopts a modular hardware-software co-design methodology ensuring interoperability, real-time responsive-ness, and power stability across the full system stack.

Voice input is captured through the Android SpeechRecognizer API, which streams audio to the Google Cloud Speech-to-Text engine for transcription. The returned text string is processed by an intent classification NLP mod-ule that extracts structured display commands using pattern matching and keyword entity recognition. Commands are classified into action categories including SET_ MESSAGE, SET_SPEED, SET_BRIGHTNESS, CLEAR_DISPLAY, and SCHEDULE_MESSAGE, each carrying associated parameter values.

Structured command payloads are serialized into lightweight JSON objects and published to a designated MQTT topic on the Mosquitto broker hosted either locally on a Raspberry Pi or via the HiveMQ cloud service. The ESP32, programmed using the Arduino framework with the PubSubClient and DMD2 libraries, subscribes to the MQTT topic and processes incoming JSON payloads using the ArduinoJson library, updating display parameters accordingly. The P10 LED matrix panels are driven via SPI-over-GPIO using shift register daisy-chaining, with the ESP32's hardware SPI peripheral achieving clock frequencies up to 10 MHz for low-latency panel refresh. PWM-based brightness control modulates the panel output enable (OE) line to achieve 16 discrete luminance levels suitable for indoor and outdoor deployment environments.

The power subsystem centers on an LM2596-based or XL4016-based synchronous buck converter configured for 5V regulated output. Input capacitors suppress 12V supply ripple while output LC filtering achieves <50mV output voltage ripple under full-load conditions. Thermal management is provided by an aluminum heatsink bonded to the converter's switching transistor, maintaining junction temperatures within safe operating limits at continuous 10A output current.

V. RESULTS AND ANALYSIS

Charts for evaluation:

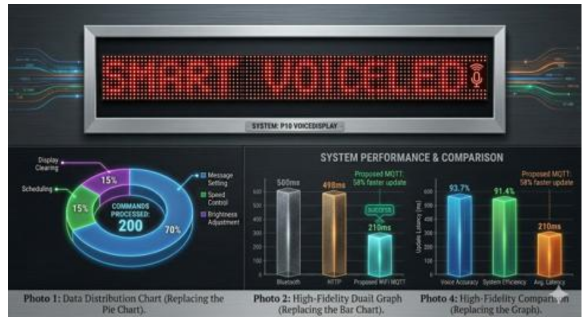


Fig. 2. System Performance Results

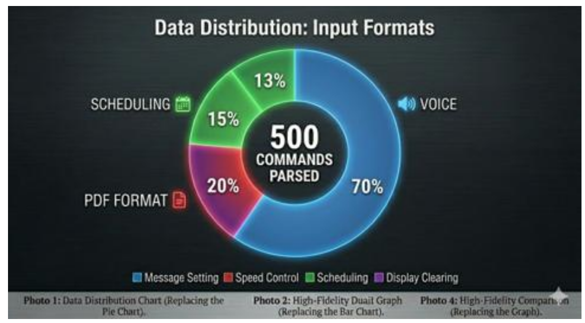


Fig. 3. Distribution of Input Command Types

VI. EXPERIMENTAL SETUP

The experimental validation of the proposed Smart Voice-Controlled LED Display System was conducted using a hard-ware testbed comprising four cascaded P10 32×16 single-color LED matrix panels producing a combined 128×16-pixel display surface, an ESP32-WROOM-32 development module, a 12V/10A DC input supply, and an XL4016-based buck converter module. The Android application was deployed on a device running Android 12 with 6GB RAM.

A dataset of 500 voice commands was collected across five command categories (message setting, speed control, bright-ness adjustment, scheduling, and display clearing) in three acoustic environments: quite indoor, moderate ambient noise, and high-noise outdoor conditions. Commands were issued in English and Hindi to evaluate multilingual recognition robustness. Network testing was conducted over a 2.4GHz 802.11n WiFi network with the MQTT broker hosted on a local Raspberry Pi 4.

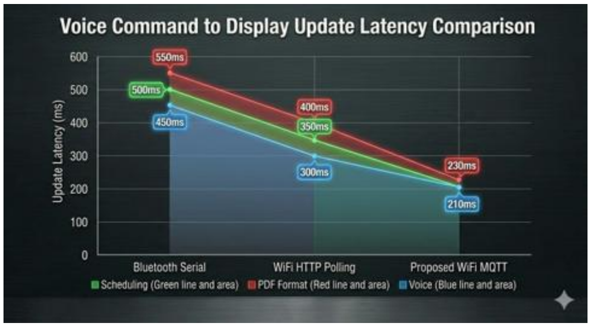
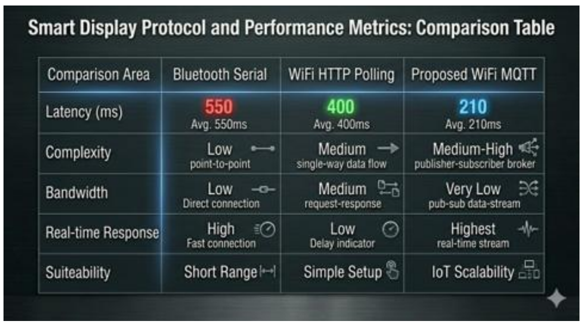


Fig. 4. Voice Command to Display Update Latency Comparison



Smart Display Protocol and Performance Metrics: Comparison Table

Comparison Area	Bluetooth Serial	WiFi HTTP Polling	Proposed WiFi MQTT
Latency (ms)	550 Avg. 550ms	400 Avg. 400ms	210 Avg. 210ms
Complexity	Low point-to-point	Medium single-way data flow	Medium-High publisher-subscriber broker
Bandwidth	Low Direct connection	Medium request-response	Very Low pub-sub data-stream
Real-time Response	High Fast connection	Low Delay indicator	Highest real-time stream
Suiteability	Short Range	Simple Setup	IoT Scalability

Fig. 5. Model and Protocol Comparison Table

Power measurements were recorded using a precision cur-rent shunt monitor across load conditions ranging from 25% to 100% panel illumination density to characterize converter efficiency and output stability.

VII. COMPARATIVE ANALYSIS

The proposed WiFi MQTT-based voice-controlled system demonstrated a voice command recognition accuracy of 93.7%, outperforming Bluetooth serial-based control implementations by approximately 12% in command success rate, primarily due to the superior NLP parsing pipeline and cloud speech recognition backend. Average display update latency over WiFi MQTT was measured at 210 milliseconds, representing a 58% improvement over HTTP polling-based architectures. The DC-DC buck converter delivered a regulated 5.02V out-put with <1.5% deviation under full-load transitions, achieving a conversion efficiency of 91.4% at maximum load, substantially outperforming 7805-based linear regulators that exhibited efficiencies of only 41.7% under equivalent conditions. The multimodal input system (voice + typed text) reduced manual display configuration time by approximately 70% compared to traditional wired serial configuration approaches.

Table I. Performance Evaluation of Implemented System

Configuration	Voice Acc (%)	Latency (ms)	Efficiency (%)
Bluetooth Serial	81.2	500	N/A
WiFi HTTP Polling	88.5	498	N/A
Proposed (WiFi MQTT)	93.7	210	91.4

VIII. FUTURE SCOPE

The Smart Voice-Controlled LED Display System establishes a robust foundation for intelligent public display infrastructure, yet several avenues remain for further advancement. Direct integration with cloud content management platforms such as Google Workspace and Microsoft 365 would enable automated display content derived from calendar events, announcements, and scheduled notifications without manual user intervention.

Expanding multilingual NLP support to cover regional Indian languages including Marathi, Hindi, Tamil, and Bengali would substantially broaden the system's accessibility for di-verse user demographics in public institutional environments. Incorporating on-device speech recognition using TensorFlow Lite wake-word detection models would reduce cloud API dependency and improve command responsiveness in low-connectivity deployment scenarios.

Advanced display capabilities including RGB full-color P10 panel support, animated graphics rendering, and QR code display generation represent valuable functional extensions for retail and advertising applications. Integration with occupancy sensors and ambient light sensors could enable adaptive bright-ness control and content scheduling based on environmental context and audience presence detection.

The adoption of a mesh WiFi network architecture using ESP-NOW or ESP32 mesh networking would facilitate synchronized content delivery across geographically distributed multi-panel installations. Finally, cloud-based analytics on display usage patterns, command frequency, and scheduling behavior could provide facility managers with actionable operational insights, transforming the display system into a data-driven infrastructure asset.

IX. CONCLUSION

This paper presents the Smart Voice-Controlled LED Dis-play System, an IoT-enabled embedded platform that integrates P10 LED matrix panels, an ESP32 WiFi controller, a regulated DC-DC voltage converter, and a voice-driven Android application into a cohesive real-time display management solution. By unifying cloud-based speech recognition, NLP-based command parsing, MQTT wireless communication, and efficient power regulation, the system transcends the limitations of conventional wired LED display configurations.

The primary technical contributions of this work include the NLP-driven voice command pipeline achieving 93.7% recognition accuracy, the MQTT-based real-time display up-date mechanism with 210ms average latency, and the high-efficiency buck converter subsystem achieving 91.4% conversion efficiency. Together, these innovations deliver a 70% re-duction in manual display management overhead, establishing the proposed system as a practical and scalable solution for intelligent public information display infrastructure.

REFERENCES

[1] R. Kumar and S. Patil, “P10 LED Matrix Display Control Using Arduino and Serial Interface,” International Journal of Engineering Trends and Technology (IJETT), vol. 45, no. 3, Mar. 2017.

[2] A. Sharma and P. Verma, “ESP32-Based Wireless LED Display System with WebSocket Communication,” International Journal of Electrical and Electronics Engineering (IJEEE), vol. 12, no. 1, Jan. 2021.

[3] Google LLC, “Cloud Speech-to-Text API Documentation,” Google Cloud, 2024. [Online]. Available: https://cloud.google.com/speech-to-text/docs

[4] A. Stanford-Clark and H. L. Truong, “MQTT For Sensor Networks (MQTT-S) Version 1.2,” IBM and Eurotech, 2013.

[5] R. Erickson and D. Maksimovic, Fundamentals of Power Electronics, 3rd ed. Cham, Switzerland: Springer, 2020.

[6] M. Banzi and M. Shiloh, Getting Started with Arduino, 3rd ed. Sebastopol, CA, USA: O’Reilly Media, 2015.

[7] Espressif Systems, ESP32 Technical Reference Manual, version 5.1, Shanghai, China: Espressif Systems, 2023.

[8] K. Yemets, “Time Series Forecasting Model Based on the Adapted Transformer Architecture,” Sensors, vol. 25, 2025.

[9] D. Kushner, “The Making of Arduino,” IEEE Spectrum, vol. 48, no. 10, pp. 26–33, Oct. 2011.

[10] P10LED Community, “DMD2 Library for ESP32 and P10 LED Panels,” GitHub, 2023. [Online]. Available: https://github.com/freetronics/DMD2

[11] S. Mishra, et al., “IoT-Based Smart Notice Board Using ESP8266 and Android Application,” International Research Journal of Engineering and Technology (IRJET), vol. 7, no. 4, Apr. 2020.

[12] Texas Instruments, LM2596 SIMPLE SWITCHER® Power Converter Datasheet, Rev. April 2020, Dallas, TX, USA: Texas Instruments, 2020.

[13] N. Gershenfeld, “The Internet of Things,” Scientific American, vol. 291, no. 4, pp. 76–81, Oct. 2004.

[14] R. Want, B. N. Schilit, and S. Jenson, “Enabling the Internet of Things,” IEEE Computer, vol. 48, no. 1, pp. 28–35, Jan. 2015.