Project Plan for T5AI-Based AI Smart Doorbell System

I. Project Overview

1.1 Project Background

The smart home security market is experiencing unprecedented growth, with video doorbells at the forefront of this revolution. According to Markets and Markets' 2024 Smart Home Security Report, the global video doorbell market is projected to reach USD 4.2 billion by 2026, with a compound annual growth rate of 18.7%. Consumer demand for intelligent, AI-powered security solutions has increased by 45% year-over-year, driven by the need for enhanced home security and convenient visitor management.

However, current market offerings present significant limitations. Dominant players like Amazon Ring and Google Nest operate within closed ecosystems, requiring proprietary cameras and monthly subscription fees ranging from $3-$10/month. According to Consumer Reports 2024, 67% of smart doorbell users express frustration with vendor lock-in, while 58% desire the ability to integrate their existing CCTV cameras with smart doorbell functionality. Additionally, 73% of users want AI-powered features such as person recognition and automated responses, which are typically only available in premium subscription tiers.

This project addresses these market gaps by leveraging the Tuya T5AI development board to create an open, AI-powered smart doorbell system. Our solution uniquely offers: - **Camera Agnostic Design**: Connect any IP camera or CCTV system - **AI Person Detection**: Cloud-based computer vision for visitor identification - **Two-Way Audio**: Real-time communication via integrated microphone and speaker - **AI Auto-Response**: Intelligent automated responses when homeowner is unavailable - **No Subscription Fees**: Self-hosted VPS solution with full user control

The project combines edge computing capabilities of the T5AI devkit with cloud-based AI processing, creating a cost-effective alternative to commercial solutions while offering superior flexibility and AI capabilities.

1.2 Project Objectives

1. **Core Function Implementation**
   - Doorbell button integration with audio alert system
   - RTSP/IP camera connection and video capture via VPS
   - AI-powered person detection and recognition
   - Real-time two-way audio communication (visitor to homeowner)
2. **AI Intelligence Features**
   - Visitor face detection and optional recognition
   - Natural language AI auto-response system
   - Event-based notification and logging
   - Multi-language support for AI responses
3. **User Experience Goals**
   - Web-based dashboard for remote monitoring and communication
   - Mobile-responsive interface for on-the-go access
   - Simple setup with 3D-printed enclosure
   - Tuya Smart Life app integration for notifications

II. Market Analysis

2.1 Market Demand (Data Support)

Residential Security Segment: According to Statista 2024, 78% of homeowners consider a video doorbell essential for home security. The average household receives 12-15 doorbell events per week, with 34% of users missing visitor interactions due to work or other commitments.

DIY Smart Home Enthusiasts: StackOverflow Developer Survey indicates that 42% of IoT enthusiasts prefer open-source or customizable solutions over proprietary systems. The maker community has grown by 28% annually, with home automation being the second-most popular project category.

Cost-Conscious Consumers: Consumer Reports shows that 61% of potential smart doorbell buyers cite subscription costs as a barrier to adoption. The average Ring Protect subscription costs $100/year, creating ongoing expenses that many find unacceptable.

AI Integration Demand: Gartner 2024 reports that 89% of smart home users want AI-powered features, but only 23% currently have access to advanced AI capabilities in their existing devices due to cost barriers.

2.2 Target User Profile

User Type	Core Needs	Usage Scenario	Willingness to Pay
Tech-savvy homeowners (30-45)	Camera flexibility, AI features, no subscriptions	Residential entry points, apartments	Medium-High (150-300 USD one-time)
DIY/Maker enthusiasts (20-40)	Customization, open-source, learning	Hobby projects, smart home integration	Medium (100-200 USD for components)
Small business owners (35-55)	Multiple camera support, visitor logging, AI screening	Office entrances, retail stores	High (300-500 USD with multi-door support)
Rental property managers (30-50)	Remote monitoring, tenant communication, cost-effective	Apartment buildings, rental properties	Medium-High (per-unit deployment)

2.3 Competitive Analysis

Feature	Amazon Ring	Google Nest	Our T5AI Solution
Camera Type Monthly Subscription	Proprietary only $3.99-$10/month	Proprietary only $6-$12/month	**Any IP/CCTV camera Free (self-hosted)**
AI Person Detection	Premium tier only	Premium tier only	**Included**

Feature	Amazon Ring	Google Nest	Our T5AI Solution
AI Auto-Response	Not available	Not available	**Full AI conversation**
Two-Way Audio	Yes	Yes	**Yes**
Custom Integration	Limited	Limited	**Full API access**
Local Processing	No	No	**Edge + Cloud hybrid**
Open Source	No	No	**Yes**

Unique Value Proposition: The only smart doorbell solution that combines any-camera compatibility, advanced AI features, and zero subscription costs in a fully customizable open-source package.

III. Technical Plan

3.1 System Architecture

3.2 Hardware Design

1. **Core Processing Unit: Tuya T5AI Development Board**
   - Dual-core ARM Cortex-A53 @ 1.5GHz
   - 8MB RAM optimized for edge processing
   - Built-in WiFi for reliable connectivity
   - GPIO pins for button and peripheral connections
   - Integrated audio codec for speaker/microphone
2. **Audio I/O Module**
   - Onboard MEMS microphone for voice capture
   - Speaker output with amplifier (GPIO39 enable)
   - Support for external speaker connection
   - Echo cancellation processing capability
3. **Doorbell Button Interface**
   - Physical momentary push button (GPIO input)
   - LED indicator for visual feedback
   - Weatherproof design considerations
   - Debounce handling in firmware
4. **Enclosure Design**
   - Custom 3D-printed case (STL files provided)
   - Weatherproof rating target: IP54
   - Wall-mount bracket included
   - Ventilation for heat dissipation
   - Dimensions: approximately 120mm x 80mm x 35mm
5. **Camera System (External)**
   - Any RTSP-compatible IP camera
   - Recommended: 1080p resolution minimum

- Night vision capability preferred
- Wide-angle lens (120°+ recommended)
- Separate power supply (camera-dependent)

3.3 Software Architecture

1. **DevKit Firmware (C/FreeRTOS)**
   - Button press detection and debouncing
   - Audio playback via Tuya audio APIs
   - Microphone capture and streaming
   - TCP client for VPS communication
   - Tuya IoT SDK integration for Smart Life notifications
   - OTA firmware update capability
2. **VPS Backend (Node.js)**
   - WebSocket server for real-time communication
   - TCP server for DevKit connection
   - RTSP camera stream capture
   - Audio transcoding and streaming
   - REST API for web dashboard
   - Authentication and security layer
3. **AI Processing Engine**
   - Computer vision for person detection (OpenCV/YOLO)
   - Face detection and optional recognition
   - Natural language processing for AI responses
   - Text-to-speech for AI-generated audio
   - Speech-to-text for visitor voice transcription
4. **Web Application (React/HTML5)**
   - Real-time video feed display
   - Two-way audio interface
   - Visitor event log and history
   - AI response configuration
   - Camera and device management
   - Mobile-responsive PWA design

3.4 Communication Protocol

Connection	Protocol	Port	Description
DevKit to VPS Control	TCP	5000	Commands, status, text messages
DevKit to VPS Audio	TCP/WebSocket	5001	Bidirectional audio stream
Camera to VPS	RTSP/UDP	554	Video stream capture
Web App to VPS	WebSocket	3000	Real-time updates and audio
VPS to AI APIs	HTTPS	443	External AI service calls

3.5 AI Features Implementation

1. **Person Detection**

- YOLO v8 or similar lightweight model
- Real-time inference on VPS
- Confidence threshold configuration
- Motion-triggered activation

2. **AI Auto-Response**
   - Integration with OpenAI GPT or local LLM
   - Customizable response personas
   - Context-aware conversation
   - Multi-turn dialogue support

3. **Audio Processing**
   - Speech-to-text transcription
   - Text-to-speech synthesis
   - Noise reduction preprocessing
   - Echo cancellation

IV. Project Implementation Plan

4.1 Phase Division

Phase 1: Research and Requirements (Week 1-2) - Analyze existing smart doorbell solutions - Define detailed functional requirements - Select camera hardware for testing - Set up development environment - Create initial system architecture document

Phase 2: Hardware Integration (Week 3-4) - Configure T5AI devkit with button input - Implement audio playback functionality - Test microphone capture and streaming - Design 3D printable enclosure - Validate WiFi connectivity and range

Phase 3: Backend Development (Week 5-7) - Implement VPS WebSocket/TCP server - Develop RTSP camera capture module - Create audio streaming pipeline - Build REST API for web dashboard - Set up authentication system

Phase 4: AI Integration (Week 8-10) - Integrate person detection model - Implement face detection capability - Connect to LLM for auto-responses - Develop speech-to-text pipeline - Create text-to-speech synthesis

Phase 5: Frontend Development (Week 11-12) - Build web dashboard interface - Implement real-time video display - Create two-way audio controls - Design visitor event log UI - Add AI configuration settings

Phase 6: Testing and Optimization (Week 13-14) - End-to-end integration testing - Performance optimization - Security audit and hardening - User acceptance testing - Bug fixes and refinements

Phase 7: Documentation and Release (Week 15-16) - Write user documentation - Create setup guides - Record demonstration videos - Prepare 3D print files - Open-source release preparation

4.2 Deliverables

Deliverable	Description	Timeline
DevKit Firmware	Complete C firmware with all features	Week 7
VPS Backend	Node.js server with AI integration	Week 12
Web Application	React-based dashboard	Week 12
3D Print Files	STL files for enclosure	Week 4
Documentation	Setup guide, API docs, user manual	Week 16
Demo Video	Full feature demonstration	Week 16

4.3 Bill of Materials (Estimated)

Component	Estimated Cost (USD)
Tuya T5AI DevKit	$25-35
External Speaker	$5-10
Doorbell Button	$2-5
3D Printing Filament	$5-10
IP Camera (1080p)	$30-80
VPS Hosting (monthly)	$5-15
Miscellaneous (wires, etc.)	$10-20
Total One-Time Cost	**$82-175**

V. Risk Assessment and Mitigation

Risk	Impact	Probability	Mitigation Strategy
Audio latency issues	High	Medium	Optimize codec, use WebRTC where possible
Camera compatibility	Medium	Medium	Document tested cameras, provide RTSP troubleshooting guide
AI processing load on VPS	Medium	Low	Implement caching, use efficient models
Network connectivity drops	High	Medium	Implement reconnection logic, local buffering
Security vulnerabilities	High	Low	Regular security audits, encrypted communications

VI. Success Metrics

1. **Functional Completeness**: All core features operational
2. **Audio Latency**: less than 500ms round-trip for two-way audio
3. **Detection Accuracy**: greater than 95% person detection rate
4. **System Uptime**: greater than 99% availability
5. **User Satisfaction**: Positive feedback from beta testers

---

VII. Conclusion

This project presents an innovative approach to smart doorbell technology by combining the flexibility of the Tuya T5AI platform with cloud-based AI processing. By offering camera-agnostic design, zero subscription costs, and advanced AI capabilities, we address the primary pain points of existing commercial solutions while providing a fully customizable, open-source alternative.

The modular architecture ensures easy expansion and customization, making it suitable for both DIY enthusiasts and developers seeking to build upon the platform. With the growing demand for intelligent, cost-effective home security solutions, this project is well-positioned to serve a significant market need.

Document Version: 1.0 Date: December 2024 Project: T5AI Smart Doorbell with AI Object Detection