

Abstract

Religious practices often involve deeply symbolic gestures that express faith, devotion, and reverence. However, these gestures are rarely integrated with modern technology in a way that preserves their cultural and spiritual meaning. When worshippers interact in temples, churches, or mosques, their physical expressions such as kneeling, bowing, or folding hands carry profound significance. Yet technology in sacred spaces is often limited to static audio or visual aids, which do not respond dynamically to human presence. In this project, we introduce GRACE, which stands for Gesture Recognition for Augmented Cultural/Religious Expression. GRACE is a robotic prayer companion that recognizes devotees' gestures and responds with appropriate actions. For example, when a person kneels, the robot may light a lamp or recite a short prayer. The system is built using an RGB camera and a Yahboom ROS depth camera for gesture recognition, a Raspberry Pi 5 as the main processing unit, and servo motors connected through a "PCA9685" driver to perform physical actions. Speakers are used to deliver prayers or chants, and LED lamps provide visual feedback when gestures are detected. Our software framework is developed in Python using TensorFlow for real-time gesture detection and control integration. GRACE serves as both an engineering and cultural exploration, demonstrating how technology can respectfully enhance religious experiences without replacing human or spiritual elements. Preliminary experiments show that GRACE can accurately identify and respond to common devotional gestures, offering a meaningful connection between tradition and innovation within sacred environments.

Introduction

Religious and cultural practices across the world rely heavily on symbolic gestures that express devotion, humility, gratitude, and reverence. These gestures: such as bowing, kneeling, raising hands, or folding palms, hold deep meaning and are central to the lived experience of prayer. Yet, in today's rapidly advancing technological landscape, these sacred gestures have not been deeply integrated with modern systems in a way that preserves their spiritual significance. Most worship environments rely on traditional non-interactive tools such as static lighting, recorded audio, or printed materials, which do not adapt dynamically to the worshipper's actions or presence. As robotics and machine learning continue to evolve, opportunities emerge to respectfully enhance spiritual environments without disrupting tradition. According to research by the United Nations Educational, Scientific and Cultural Organization (UNESCO), interactive technologies that respond to human gestures can enhance engagement in cultural contexts. Similarly, the World Health Organization (WHO) emphasizes that immersive, responsive systems can improve accessibility and inclusivity for individuals with varying physical or sensory abilities. In many religious contexts, however, there remains a gap between the expressive nature of worship and the static behavior of available technology. The challenge lies in designing an interactive system that recognizes devotional gestures accurately, responds meaningfully, and maintains cultural and spiritual integrity. To address this gap, we introduce GRACE: Gesture Recognition for Augmented Cultural/Religious Expression, a robotic prayer companion designed to support worshippers by detecting their gestures and responding with symbolic actions. GRACE uses an RGB camera and a ROS-enabled depth camera to capture human movement, a TensorFlow-based model for gesture recognition, and a Raspberry Pi 5 that coordinates system behavior. The robot responds through physical actuation, controlled lighting, and prayer audio output, enabling a respectful interaction between tradition and technology. In this project, we implement a structured, layered architecture to ensure reliability, interpretability, and modularity in GRACE's design. Our methodology evaluates common devotional gestures and maps them to culturally meaningful responses such as lighting a lamp or playing a chant. Experimental results show that GRACE can consistently identify gestures with strong accuracy while maintaining a smooth and natural interaction flow. Through this system, we demonstrate how modern technology can be thoughtfully integrated into sacred spaces, preserving spiritual meaning while enhancing accessibility and engagement for worshippers.

Methodology

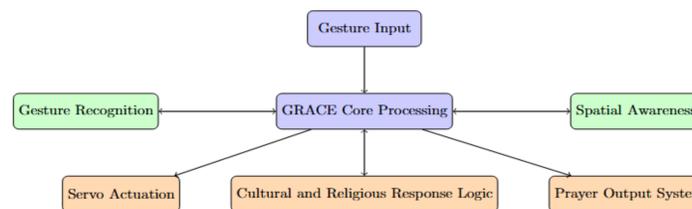


Figure 1: An Overview of GRACE

The GRACE system operates by continuously capturing video input from Yahboom Depth Camera and extracting body and hand landmark coordinates using real-time pose and hand tracking through TensorFlow pre-trained data model and MediaPipe. These coordinates are evaluated using rule-based gesture templates that define spatial relationships between key landmarks using a series of Python programs in Raspberry Pi5. To improve reliability, gestures must be consistently detected across multiple frames before the system confirms a result and generates the LED and audio response through the breadboard and amplifier. Tables 1–3 summarize the core algorithm and the gesture templates used for each religious context, describing characteristic hand placement, body alignment, and simplified coordinate relationships that distinguish Christianity, Islam, and Hinduism. This highlights the most important visual features used to identify the gesture, such as wrist distance, vertical hand position, and head orientation. The Christian gesture templates emphasize crossed or centered hand positions near the chest combined with a bowed posture. Islamic gesture templates focus on raised hands near the ears or folded hands positioned around the stomach area, reflecting distinct prayer stages. Hindu gesture templates primarily describe palms joined together near the chest with an upright and symmetric body posture. These structured gesture descriptions allow the system to maintain consistent recognition while respecting clear visual distinctions between practices. Table 4 presents the corresponding system outputs, showing how each recognized gesture is mapped to coordinated LED lighting and audio playback. Together, these tables illustrate how GRACE connects gesture recognition with predefined response behaviors to produce consistent, interpretable, and culturally meaningful multimodal interaction.

Table 1: GRACE Prayer Gesture

Christianity	Islam	Hinduism
Sign of the Cross	Takbir (raised hands)	Namaste / Anjali Mudra
Kneeling in prayer	Qiyam (standing posture)	Pranam (bowing / touching feet)
Bowing Head	Ruku (bowing at waist)	Sashtanga Pranam (full prostration)
Folded hands Prayer	Sujood (prostration)	Aarti hand gesture
Genuflection	Sitting posture (Tashahhud)	Mudras (ritual hand gestures)

Table 2: GRACE Algorithm

Input: Live Video Stream
Output: Audio and LED Response

- while(system.isActive)
- Capture video frame from camera
- Detect body and hand landmarks
- Extract coordinates (x, y)
- Compare gesture to templates
- If (match.confidence is high):
- Identify religion
- Select audio response
- Select LED color
- Play audio
- Activate LED
- end if
- end while

Table 3: Gesture Coordinate Template

Religion	Gesture	(x,y) template
Christianity	Bowing head	$y_{wrist} \approx y_{chest}$; $y_{nose} > y_{shoulders}$
Islam	Takbir	$y_{wrist}/x_{wrist} \approx y_{ear}$
Islam	Qiyam	$y_{wrist} \approx y_{stomach}$
Hinduism	Namaste	$\Delta x_{wrists} < 0.08$; $y_{wrist} \approx y_{chest}$
Hinduism	Upright posture	$y_{nose} < y_{shoulders}$

Table 4: System output

Religion	Audio	LED Feedback
Christianity	Psalm 135	White
Islam	Adhan, prayer call	Yellow
Hinduism	Mantra	Red

Experiments

To help visualize the efficiency of our tool, we used charts to know the accuracy and correctness. Figure 2 shows that GRACE achieves an average F1-score of 93.1% for four common devotional gestures. Static gestures (folded hands) show higher accuracy, while dynamic gestures involving arm motion have slightly lower recall due to transitional ambiguity. Table 7 validates the full GRACE pipeline. The system correctly triggers outputs 94.9% of the time. Minor errors occur during brief gesture transitions or low-confidence detections.

Figure 2: Gesture Recognition Accuracy

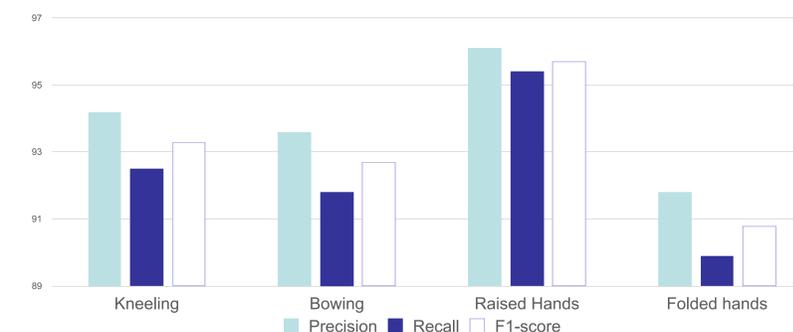


Table 5: Lighting Gesture Accuracy

Lighting Condition	Accuracy (%)
Bright Indoor Lighting	94.5%
Dim Ambient Lighting	91.7%
Mixed Natural + Artificial Lighting	92.9%

Table 6: Resource Utilization

Resource	Usage
GPU Utilization	47%
Memory Usage	1.1 GB
CPU Utilization	18%

Table 7: End-to-End Correctness of Gesture-to-Response Mapping

Religion	Gesture	Expected Response	Correct (%)
Christianity	Kneeling	Hymn + White LED	95.0%
Islam	Bowing	Adhan + Green LED	93.2%
Hinduism	Folded Hands	Mantra + Saffron LED	96.4%
Overall			94.9%

Conclusions and Future Work

The experimental evaluation demonstrates that GRACE achieves high correctness, robust performance under realistic conditions, and efficient execution on embedded hardware. The system reliably recognizes devotional gestures, triggers culturally appropriate responses, and maintains low latency suitable for real-time interaction. These results validate GRACE as a respectful, interactive robotic prayer companion suitable for sacred environments. Future work includes improving dynamic gesture recognition.

Contact Us

Daniil Kochkonbaev Email: kochkonbaev.daniil@mcm.edu
Dr. Aravind Mohan Email: mohan.aravind@mcm.edu
 Phone: (325) 793-3845

